



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> :  C12Q 1/68		A2	(11) International Publication Number: <b>WO 00/12760</b>																					
			(43) International Publication Date: 9 March 2000 (09.03.00)																					
<p>(21) International Application Number: PCT/US99/19768</p> <p>(22) International Filing Date: 27 August 1999 (27.08.99)</p> <p>(30) Priority Data:</p> <table> <tr><td>09/141,825</td><td>28 August 1998 (28.08.98)</td><td>US</td></tr> <tr><td>09/172,711</td><td>13 October 1998 (13.10.98)</td><td>US</td></tr> <tr><td>09/172,108</td><td>13 October 1998 (13.10.98)</td><td>US</td></tr> </table> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</p> <table> <tr><td>US</td><td>09/141,825 (CIP)</td></tr> <tr><td>Filed on</td><td>28 August 1998 (28.08.98)</td></tr> <tr><td>US</td><td>09/172,711 (CIP)</td></tr> <tr><td>Filed on</td><td>13 October 1998 (13.10.98)</td></tr> <tr><td>US</td><td>09/172,108 (CIP)</td></tr> <tr><td>Filed on</td><td>13 October 1998 (13.10.98)</td></tr> </table> <p>(71) Applicant (<i>for all designated States except US</i>): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>for US only</i>): CUNNINGHAM, Mary, Jane [US/US]; 1204 Manet Drive, Sunnyvale, CA 94087</p>			09/141,825	28 August 1998 (28.08.98)	US	09/172,711	13 October 1998 (13.10.98)	US	09/172,108	13 October 1998 (13.10.98)	US	US	09/141,825 (CIP)	Filed on	28 August 1998 (28.08.98)	US	09/172,711 (CIP)	Filed on	13 October 1998 (13.10.98)	US	09/172,108 (CIP)	Filed on	13 October 1998 (13.10.98)	<p>(US). ZWEIGER, Gary, B. [US/US]; 2700 Del Medio Court, #124, Mountain View, CA 95023 (US). PANZER, Scott, R. [US/US]; 965 East El Camino, #621, Sunnyvale, CA 94087 (US). SEILHAMER, Jeffrey, J. [US/US]; 12555 La Cresta, Los Altos, CA 94022 (US).</p> <p>(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p>
09/141,825	28 August 1998 (28.08.98)	US																						
09/172,711	13 October 1998 (13.10.98)	US																						
09/172,108	13 October 1998 (13.10.98)	US																						
US	09/141,825 (CIP)																							
Filed on	28 August 1998 (28.08.98)																							
US	09/172,711 (CIP)																							
Filed on	13 October 1998 (13.10.98)																							
US	09/172,108 (CIP)																							
Filed on	13 October 1998 (13.10.98)																							
<p>(54) Title: TOXICOLOGICAL RESPONSE MARKERS</p> <p>(57) Abstract</p> <p>The present invention relates to a composition comprising a plurality of nucleic acid molecules. The composition can be used as hybridizable array elements in a microarray. The present invention also relates to methods for screening compounds and therapeutics for metabolic responses indicative of a toxic compound.</p>																								

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## TOXICOLOGICAL RESPONSE MARKERS

This application is filled under the Patent Cooperation Treaty and claims the benefits of U.S. Nonprovisional Application No. 09/141,825, our Docket No. PA-0010 US, filed on August 5 28, 1998, U.S. Nonprovisional Application No. 09/172,711, our Docket No. PA-0011 US, filed on October 13, 1998, and U.S. Nonprovisional Application No. 09/172,108, our Docket No. PA-0012 US, filed on October 13, 1998.

### TECHNICAL FIELD

The present invention relates to compositions and methods for use in detecting metabolic 10 and toxicological responses.

### BACKGROUND ART

Toxicity testing is a mandatory and time-consuming part of the pharmaceutical drug development pipeline. A more rapid screen to determine the effects upon metabolism and to detect toxicity of lead drug candidates may be the use of gene expression microarrays. For example, 15 microarrays of various kinds may be produced using full length genes or gene fragments. These arrays can then be used to test samples treated with the drug candidates to elucidate the gene expression pattern associated with drug treatment. This gene pattern can be compared with gene expression patterns associated with compounds which produce known toxicological and metabolic responses.

20 Benzo(a)pyrene is a known rodent and likely human carcinogen and is the prototype of a class of compounds, the polycyclic aromatic hydrocarbons (PAH). It is metabolized by several forms of cytochrome P450 (P450 isozymes) and associated enzymes to form both activated and detoxified metabolites. The ultimate metabolites are the bay-region diol epoxide, benzo(a)pyrene-7,8-diol-9,10-epoxide (BPDE) and the K-region diol epoxide, 9-hydroxy benzo(a)pyrene-4,5-oxide, both of which induce formation of DNA adducts. DNA adducts have been shown to persist 25 in rat liver up to 56 days following treatment with benzo(a)pyrene at a dose of 10 mg/kg body weight three times per week for two weeks (Qu and Stacey (1996) Carcinogenesis 17:53-59).

Acetaminophen is a widely-used analgesic. It is metabolized by specific cytochrome P450 isozymes with the majority of the drug undergoing detoxification by glucuronic acid, sulfate and 30 glutathione conjugation pathways. However, at supratherapeutic doses, acetaminophen is metabolized to an active intermediate, *N*-acetyl-*p*-benzoquinone imine (NAPQI) which can cause hepatic and renal failure. NAPQI then binds to sulfhydryl groups of proteins causing their inactivation and leading to subsequent cell death (Kroger *et al.* (1997) Gen. Pharmacol. 28:257-263).

35 Clofibrate is an hypolipidemic drug which lowers elevated levels of serum triglycerides.

In rodents, chronic treatment produces hepatomegaly and an increase in hepatic peroxisomes. Clofibrate has been shown to increase levels of cytochrome P450 4A and reduce the levels of P450 4F. It is also involved in transcription of  $\beta$ -oxidation genes as well as induction of peroxisome proliferator (PP) activated receptors (Kawashima *et al.* (1997) Arch. Biochem. Biophys. 347:148-154). Peroxisome proliferation that is induced by both clofibrate and the chemically-related compound fenofibrate is mediated by a common inhibitory effect on mitochondrial membrane depolarization (Zhou and Wallace (1999) Toxicol. Sci. 48:82-89).

The present invention provides compositions and methods for the screening of compounds for metabolic and toxicological responses.

10

#### DISCLOSURE OF INVENTION

15

The invention provides nucleic acid molecules whose transcript levels are modulated in a sample during a metabolic response to a toxic compound. The invention also provides nucleic acid molecules whose transcript levels are upregulated in a sample during a metabolic response to a toxic compound. The invention also provides nucleic acid molecules whose transcript levels are downregulated in a sample during a metabolic response to a toxic compound. Upregulation or downregulation is at least 2 fold, more preferably at least 2.5 fold, most preferably at least 3 fold. The metabolic response to a toxic compound may be a toxicological response.

20

25

30

In another aspect, the invention provides a method for screening a compound for a metabolic response to a test compound or molecule. The method comprises treating a tissue with a known toxic compound, determining levels of a plurality of nucleic acid molecules, selecting from the plurality of nucleic acid molecules those nucleic acid molecules that have levels modulated in samples treated with known toxic compounds when compared with untreated samples. Some of the transcript levels may be upregulated by a toxic compound, others may be downregulated by a toxic compound, and still others may be upregulated with one known toxic compound and be downregulated with another known toxic compound. The selected nucleic acid molecules which are upregulated and downregulated by a known toxic compound are arrayed upon a substrate. The method further comprises determining levels of nucleic acid molecules in a sample after the sample is treated with a compound. Levels of nucleic acid molecules in a sample so treated are then compared with the plurality of the arrayed nucleic acid molecules to identify which sample nucleic acid molecules are upregulated and downregulated by the compound.

35

Preferably, the comparing comprises contacting the arrayed nucleic acid molecules with the sample nucleic acid molecules under conditions effective to form hybridization complexes between the arrayed nucleic acid molecules and the sample nucleic acid molecules; and detecting the presence or absence of the hybridization complexes. In this context, similarity may mean that at least 1, preferably at least 5, more preferably at least 10, of the upregulated arrayed nucleic acid

molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a greater extent than would the probes derived from a sample not treated with the test compound or a known toxic compound. Similarity may also mean that at least 1, preferably at least 3, of the downregulated arrayed nucleic acid molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a lesser extent than would the sample nucleic acid molecules of a sample not treated with the test compound or a known toxic compound.

- Preferred tissues are selected from the group consisting of liver, kidney, brain, spleen, pancreas and lung. Preferred toxic compounds are selected from the group consisting of
- hypolipidemic drugs, n-alkylcarboxylic acids, n-alkylcarboxylic acid precursors, azole antifungal compounds, leukotriene D4 antagonists, herbicides, pesticides, phthalate esters, phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol and their corresponding metabolites, acetaminophen and its corresponding metabolites, benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like.
- The arrayed nucleic acid molecules comprise fragments of messenger RNA transcripts of genes that are up-regulated or down-regulated at least 2-fold, preferably at least 2.5-fold, more preferably at least 3-fold, in samples treated with known toxic compounds when compared with untreated samples. Preferred arrayed nucleic acid molecules are selected from the group consisting of SEQ ID NOs:1-117, or fragments thereof, some of whose levels are upregulated and others of whose levels are downregulated. More preferable are SEQ ID NOs: 3, 9, 10, 13, 19, 26, 31, 33, 35, 36, 37, 39, 42, 57, 67, 78, 81, 82, 94, and 98 which are upregulated, and SEQ ID NOs: 43, 49, 50, 52, 53, 54, 55, 56, 59, 61, 63, 68, 71, 74, 85, 87, 90, 95, 102, 103, 105, and 115 which are downregulated. Most preferable are SEQ ID NOs: 31, 33, 35, 36, 39, 52, 53, 54, 55, 63, 74, 81, 90, 94, and 95. In one embodiment, the polynucleotide targets are hybridizable array elements of a microarray.

Alternatively, the invention provides methods for screening a test compound or molecule for a metabolic response or for screening a sample for a metabolic response to a test compound or molecule.

Alternatively, the invention provides methods for screening a test compound or molecule for a previously unknown metabolic response or for screening a sample for a previously unknown metabolic response to a test compound or molecule.

In another aspect, the invention provides methods for preventing a toxicological response by administering complementary nucleotide molecules against one or more selected upregulated nucleic acid molecules or a ribozyme that specifically cleaves such molecules. Alternatively, a toxicological response may be prevented by administering sense nucleotide molecules for one or

more selected downregulated nucleic acid molecules.

In yet another aspect, the invention provides methods for preventing a toxicological response by administering an agonist which initiates transcription of a gene comprising a downregulated nucleic acid molecule of the invention. Alternatively, a toxicological response may 5 be prevented by administering an antagonist which prevents transcription of a gene comprising an upregulated nucleic acid molecule of the invention.

The invention also provides a substantially purified mammalian protein or a portion thereof. The invention further provides isolated and purified proteins encoded by the nucleic acid molecules of SEQ ID NOs:1-117. Additionally, the invention provides a pharmaceutical 10 composition comprising a substantially purified mammalian protein or a portion thereof in conjunction with a pharmaceutical carrier.

The invention further provides a method for using at least a portion of the proteins encoded by SEQ ID NOs:1-117 to produce antibodies. The invention also provides a method for using a protein or a portion thereof to screen a library of molecules to identify at least one ligand 15 which specifically binds the protein, the method comprising combining the protein with the library of molecules under conditions allowing specific binding, and detecting specific binding, thereby identifying a ligand which specifically binds the protein. Such libraries include DNA and RNA molecules, peptides, agonists, antagonists, antibodies, immunoglobulins, drug compounds, pharmaceutical agents, and other ligands. In one aspect, the ligand identified using the method 20 modulates the activity of the mammalian protein. In an analogous method, the protein or a portion thereof is used to purify a ligand. The method involves combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and separating the protein from the ligand to obtain purified ligand.

25 The invention further provides a method for inserting a marker gene into the genomic DNA of an animal to disrupt the expression of the natural nucleic acid molecule. The invention also provides a method for using the nucleic acid molecule to produce an animal model system, the method comprising constructing a vector containing the nucleic acid molecule; introducing the vector into a totipotent embryonic stem cell; selecting an embryonic stem cell with the vector 30 integrated into genomic DNA; microinjecting the selected cell into a blastocyst, thereby forming a chimeric blastocyst; transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric animal containing at least one additional copy of nucleic acid molecule in its germ line; and breeding the chimeric animal to generate a homozygous animal model system.

**BRIEF DESCRIPTION OF THE SEQUENCE LISTING**

A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

The Sequence Listing contains the nucleic acid sequence of exemplary nucleic acid molecules of the invention, SEQ ID NOS:1-117.

**MODES FOR CARRYING OUT THE INVENTION****10 Definitions**

"Sample" is used in its broadest sense. A sample containing nucleic acid molecules may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; a cell; a biological tissue or isolated fragment thereof, for example, a needle biopsy; a fingerprint or tissue print; natural or synthetic fibres; in a solution; in a liquid suspension; in a gaseous suspension; in an aerosol; and the like.

"Plurality" refers preferably to a group of one or more members, preferably to a group of at least about 10, and more preferably to a group of at least about 100 members, and even more preferably a group of 10,000 members.

"Substrate" refers to a rigid or semi-rigid support to which nucleic acid molecules or proteins are bound and includes membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, capillaries or other tubing, plates, polymers, and microparticles with a variety of surface forms including wells, trenches, pins, channels and pores.

"Modulates" refers to a change in activity (biological, chemical, or immunological) or lifespan resulting from specific binding between a molecule and either a nucleic acid molecule or a protein

"Microarray" refers to an ordered arrangement of hybridizable array elements on a substrate. The array elements are arranged so that there are preferably at least ten or more different array elements, more preferably at least 100 array elements, even more preferably at least 1000 array elements, and most preferably 10,000. Furthermore, the hybridization signal from each of the array elements is individually distinguishable. In a preferred embodiment, the array elements comprise nucleic acid molecules.

"Nucleic acid molecule" refers to a nucleic acid, oligonucleotide, nucleotide, polynucleotide or any fragment thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein or other

materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA). "Oligonucleotide" is substantially equivalent to the terms amplimer, primer, oligomer, element, target, and probe and is preferably single stranded.

"Protein" refers to an amino acid sequence, oligopeptide, peptide, polypeptide or portions thereof whether naturally occurring or synthetic.

"Up-regulated" refers to a nucleic acid molecule whose levels increased in a treated sample compared with the nucleic acid molecule in an untreated sample.

"Down-regulated" refers to nucleic acid molecule whose levels decreased in a treated sample compared with the nucleic acid molecule in an untreated sample.

"Toxic compound" or "toxic agent" is any compound, molecule, or agent that elicits a biochemical, metabolic, and physiological response in an individual or animal, such as i) DNA damage, ii) cell damage, iii) organ damage or cell death, or iv) clinical morbidity or mortality.

"Toxicological response" refers to a biochemical, metabolic, and physiological response in an individual, animal, or test system which has been exposed to a toxic compound or toxic agent.

"Fragment" refers to an Incyte clone or any part of a nucleic acid molecule which retains a usable, functional characteristic. Useful fragments include oligonucleotides and polynucleotides which may be used in hybridization or amplification technologies or in regulation of replication, transcription or translation. Exemplary fragments are the first twenty consecutive nucleotides of SEQ ID NOs:1-117.

"Hybridization complex" refers to a complex between two nucleic acid molecules by virtue of the formation of hydrogen bonds between purines and pyrimidines.

"Ligand" refers to any molecule, agent, or compound which will bind specifically to a complementary site on a nucleic acid molecule or protein. Such ligands stabilize or modulate the activity of nucleic acid molecules or proteins of the invention and may be composed of at least one of the following: inorganic and organic substances including nucleic acids, proteins, carbohydrates, fats, and lipids.

"Substantially purified" refers to nucleic acid molecules or proteins that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free, from other components with which they are naturally associated.

### The Invention

The present invention provides a composition and method of using the composition for screening test compounds and molecules for toxicological responses. Additionally the invention provides methods for characterizing the toxicological responses of a sample to a test compound or molecule. In particular, the present invention provides a composition comprising a plurality of

nucleic acid molecules derived from human cDNA libraries, monkey cDNA libraries, mouse cDNA libraries, normal rat liver cDNA libraries, normalized rat liver cDNA libraries and prehybridized rat liver cDNA libraries and rat kidney cDNA libraries in a test system. The nucleic acid molecules have been further selected for exhibiting up-regulated or down-regulated gene expression in rat livers when the rats have been exposed to a known hepatotoxin, including a peroxisome proliferator (PP), acetaminophen or one of its corresponding metabolites, and a polycyclic aromatic hydrocarbon (PAH).

PPs include hypolipidemic drugs, such as clofibrate, fenofibrate, clofenic acid, nafenopin, gemfibrozil, ciprofibrate, bezafibrate, halofenate, simfibrate, benzofibrate, etofibrate, WY-14,643, and the like; n-alkylcarboxylic acids, such as trichloroacetic acid, valproic acid, hexanoic acid, and the like; n-alkylcarboxylic acid precursors, such as trichloroethylene, tetrachloroethylene, and the like; azole antifungal compounds, such as bifenazole, and the like; leukotriene D4 antagonists; herbicides; pesticides; phthalate esters, such as di-[2-ethylhexyl] phthalate, mono-[2-ethylhexyl] phthalate, and the like; and natural chemicals, such as phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol, and the like. In a preferred embodiment the toxic compound is clofibrate, or one of its corresponding metabolites. In another preferred embodiment the toxic compound is fenofibrate, or one of its corresponding metabolites.

PAHs include compounds such as benzo(a)pyrene, 3-methylcholanthrene, benzo(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like. In a preferred embodiment the toxic compound is benzo(a)pyrene, or one of its corresponding metabolites.

SEQ ID NOs:1-117 were identified by their pattern of at least two-fold up-regulation or down-regulation following hybridization with sample nucleic acid molecules from treated rat liver tissue. These and other nucleic acid molecules can be immobilized on a substrate as hybridizable array elements in a microarray format. The microarray may be used to characterize gene expression patterns associated with novel compounds to elucidate any metabolic responses or to monitor the effects of treatments during clinical therapy where metabolic responses to toxic compounds may be expected.

When the nucleic acid molecules are employed as hybridizable array elements in a microarray, the array elements are organized in an ordered fashion so that each element is present at a specified location on the substrate. Because the array elements are at specified locations on the substrate, the hybridization patterns and intensities (which together create a unique expression profile) can be interpreted in terms of expression levels of particular genes and can be correlated with a toxicological response associated with a test compound or molecule.

Furthermore, the present invention provides methods for screening test compounds and/or

molecules for potential toxicological responses and for screening a sample's toxicological response to a particular test compound or molecule. Briefly, these methods entail treating a sample with the test compound or molecule to elicit a change in gene expression patterns comprising the expression of a plurality of sample nucleic acid molecules. Nucleic acid molecules  
5 are selected by identifying those levels of expressed nucleic acid molecules in rat liver or kidney which are up-regulated or down-regulated at least 2-fold, more preferably at least 2.5-fold, most preferably at least 3-fold, when treated with a known toxic compound. The nucleic acid molecules are arrayed on a substrate. Then, the arrayed nucleic acid molecules and sample nucleic acid molecules are combined under conditions effective to form hybridization complexes which may be  
10 detected by methods well known in the art. Detection of higher or lower levels of such hybridization complexes compared with hybridization complexes derived from samples treated with a compound that is known not to induce a toxicological response correlates with a toxicological response to a test compound or a toxicological response to a molecule.

Complementary DNA libraries

15 Molecules are identified that reflect all or most of the genes that are expressed in rat tissue. Molecules may be identified by isolating clones derived from several types of rat cDNA libraries, including normal rat cDNA libraries, normalized rat cDNA libraries and prehybridized rat cDNA libraries. Clone inserts derived from these clones may be partially sequenced to generate expressed sequence tags (ESTs).

20 In one embodiment, two collections of ESTs are identified and sequenced. A first collection of ESTs (the originator molecules) are derived from rat liver and kidney and from the cDNA libraries presented in the Examples. A second collection includes ESTs derived from other rat cDNA libraries available in the ZOSEQ database (Incyte Pharmaceuticals, Inc., Palo Alto CA).

25 The two collections of ESTs are clustered electronically to form master clusters of ESTs. Master clusters are formed by identifying overlapping EST molecules and assembling these ESTs. A nucleic acid fragment assembly tool, such as the Phrap tool (Phil Green, University of Washington) and the GELVIEW fragment assembly system (GCG, Madison WI), can be used for this purpose. The minimum number of clones which constitute a cluster is two. In another  
30 embodiment, a collection of human genes known to be expressed in response to toxic agents are used to select representative ESTs from the 113 rat cDNA libraries. The master cluster process is repeated for these molecules.

After assembling the clustered consensus nucleic acid sequences, a representative 5' clone is nominated from each master cluster. The most 5' clone is preferred because it is most likely to  
35 contain the complete gene. The nomination process is described in greater detail in "Relational

Database and System for Storing Information Relating to Biomolecular Sequences and Reagents", USSN 09/034,807, filed March 4, 1998, herein incorporated in its entirety by reference. The EST molecules are used as array elements on a microarray.

Selection of arrayed nucleic acid molecules

5 Samples are treated, preferably at subchronic doses, with one or more known toxic compounds over a defined time course. Preferably, the agents are peroxisomal proliferators (PPs), acetaminophen or one of its corresponding metabolites, and polycyclic aromatic hydrocarbons (PAHs).

The gene expression patterns derived from such treated biological samples can be  
10 compared with the gene expression patterns derived from untreated biological samples to identify nucleic acid molecules whose expression is either up-regulated or down-regulated due to the response to the toxic compounds. These molecules may then be employed as array elements alone or in combination with other array element molecules. Such a microarray is particularly useful to detect and characterize gene expression patterns associated with known toxic compounds. Such  
15 gene expression patterns can then be used for comparison to identify other compounds which also elicit a metabolic response to a toxic compound.

The arrayed nucleic acid molecules can be manipulated to optimize their performance in hybridization. To optimize hybridization, the arrayed nucleic acid molecules are examined using a computer algorithm to identify portions of genes without potential secondary structure. Such  
20 computer algorithms are well known in the art and are part of OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or LASERGENE software (DNASTAR, Madison WI). These programs can search within nucleic acid molecule sequences to identify stem loop structures and tandem repeats and to analyze G + C content of the sequence (those molecules with a G + C content greater than 60% are excluded). Alternatively, the arrayed nucleic acid molecules  
25 can be optimized by trial and error. Experiments can be performed to determine whether sample nucleic acid molecules and complementary arrayed nucleic acid molecules hybridize optimally under experimental conditions.

The arrayed nucleic acid molecules can be any RNA-like or DNA-like material, such as mRNAs, cDNAs, genomic DNA, peptide nucleic acids, branched DNAs and the like. The arrayed  
30 nucleic acid molecules can be in sense or antisense orientations.

In one embodiment, the arrayed nucleic acid molecules are cDNAs. The size of the DNA sequence of interest may vary, and is preferably from 50 to 10,000 nucleotides, more preferably from 150 to 3,500 nucleotides. In a second embodiment, the nucleic acid molecules are vector DNAs. In this case the size of the DNA sequence of interest, i.e., the insert sequence, may vary  
35 from about 50 to 10,000 nucleotides, more preferably from about 150 to 3,500 nucleotides.

The nucleic acid molecule sequences of the Sequence Listing have been prepared by current, state-of-the-art, automated methods and, as such, may contain occasional sequencing errors and unidentified nucleotides. Nucleotide analogues can be incorporated into the nucleic acid molecules by methods well known in the art. The only requirement is that the incorporated 5 nucleotide analogues must serve to base pair with sample nucleic acid molecules. For example, certain guanine nucleotides can be substituted with hypoxanthine which base pairs with cytosine residues. However, these base pairs are less stable than those between guanine and cytosine. Alternatively, adenine nucleotides can be substituted with 2,6-diaminopurine which can form 10 stronger base pairs than those between adenine and thymidine. Additionally, the nucleic acid molecules can include nucleotides that have been derivatized chemically or enzymatically. Typical modifications include derivatization with acyl, alkyl, aryl or amino groups.

The nucleic acid molecules can be immobilized on a substrate via chemical bonding. Furthermore, the molecules do not have to be directly bound to the substrate, but rather can be bound to the substrate through a linker group. The linker groups are typically about 6 to 50 atoms 15 long to provide exposure to the bound nucleic acid molecule. Preferred linker groups include ethylene glycol oligomers, diamines, diacids and the like. Reactive groups on the substrate surface react with one of the terminal portions of the linker to bind the linker to the substrate. The other terminal portion of the linker is then functionalized for binding the nucleic acid molecule. Preferred substrates are any suitable rigid or semirigid support, including membranes, filters, 20 chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which the arrayed nucleic acid molecules are bound.

The samples can be any sample comprising sample nucleic acid molecules and obtained from any bodily fluid (blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or 25 other tissue preparations. The samples can be derived from any species, but preferably from eukaryotic species, and more preferably from mammalian species such as rat and human.

DNA or RNA can be isolated from the sample according to any of a number of methods well known to those of skill in the art. For example, methods of purification of nucleic acids are described in Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation, P. Tijssen, ed. Elsevier (1993). In one preferred embodiment, total RNA is isolated using the TRIZOL total RNA isolation reagent (Life Technologies, Gaithersburg MD) and mRNA is isolated using oligo d(T) column chromatography or glass beads. When sample nucleic acid molecules are amplified it is desirable to amplify the sample nucleic acid molecules and maintain the relative abundances of the original 35 sample, including low abundance transcripts. RNA can be amplified in vitro, in situ or in vivo.

(See Eberwine US Patent No. 5,514,545).

It is also advantageous to include controls within the sample to assure that amplification and labeling procedures do not change the true distribution of nucleic acid molecules in the sample. For this purpose, a sample is spiked with an amount of control nucleic acid molecules predetermined to be detectable upon hybridization to its complementary arrayed nucleic acid molecule and the composition of nucleic acid molecules includes reference nucleic acid molecules which specifically hybridize with the control arrayed nucleic acid molecules. After hybridization and processing, the hybridization signals obtained should reflect accurately the amounts of control arrayed nucleic acid molecules added to the sample.

Prior to hybridization, it may be desirable to fragment the sample nucleic acid molecules. Fragmentation improves hybridization by minimizing secondary structure and cross-hybridization to other sample nucleic acid molecules in the sample or noncomplementary nucleic acid molecules. Fragmentation can be performed by mechanical or chemical means.

#### Labeling

The sample nucleic acid molecules may be labeled with one or more labeling moieties to allow for detection of hybridized arrayed/sample nucleic acid molecule complexes. The labeling moieties can include compositions that can be detected by spectroscopic, photochemical, biochemical, bioelectronic, immunochemical, electrical, optical or chemical means. The labeling moieties include radioisotopes, such as  $^{32}\text{P}$ ,  $^{33}\text{P}$  or  $^{35}\text{S}$ , chemiluminescent compounds, labeled binding proteins, heavy metal atoms, spectroscopic markers, such as fluorescent markers and dyes, magnetic labels, linked enzymes, mass spectrometry tags, spin labels, electron transfer donors and acceptors, and the like. Preferred fluorescent markers include Cy3 and Cy5 fluorophores (Amersham Pharmacia Biotech, Piscataway NJ).

#### Hybridization

The nucleic acid molecule sequence of SEQ ID NOs:1-117 and fragments thereof can be used in various hybridization technologies for various purposes in a test system. Hybridization probes may be designed or derived from SEQ ID NOs:1-117. Such probes may be made from a highly specific region such as the 5' regulatory region or from a conserved motif, and used in protocols to identify naturally occurring sequences encoding the mammalian protein, allelic variants, or related sequences, and should preferably have at least 50% sequence identity to any of the protein sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NOs:1-117 or from genomic sequences including promoters, enhancers, and introns of the mammalian gene. Hybridization or PCR probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of the labeled nucleotide. A vector containing the nucleic acid sequence may be used to

produce an mRNA probe in vitro by addition of an RNA polymerase and labeled nucleic acid molecules. These procedures may be conducted using commercially available kits such as those provided by Amersham Pharmacia Biotech.

The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature. Hybridization can be performed at low stringency with buffers, such as 5 x SSC with 1% sodium dodecyl sulfate (SDS) at 60°C, which permits the formation of a hybridization complex between nucleotide sequences that contain some mismatches. Subsequent washes are performed at higher stringency with buffers such as 0.2 x SSC with 0.1% SDS at either 45°C (medium stringency) or 68°C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acid sequences are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of other detergents such as Sarkosyl or Triton X-100 and a blocking agent such as salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (*supra*) and Sambrook *et al.* (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.

Hybridization specificity can be evaluated by comparing the hybridization of specificity-control nucleic acid molecules to specificity-control sample nucleic acid molecules that are added to a sample in a known amount. The specificity-control arrayed nucleic acid molecules may have one or more sequence mismatches compared with the corresponding arrayed nucleic acid molecules. In this manner, whether only complementary arrayed nucleic acid molecules are hybridizing to the sample nucleic acid molecules or whether mismatched hybrid duplexes are forming is determined.

Hybridization reactions can be performed in absolute or differential hybridization formats. In the absolute hybridization format, nucleic acid molecules from one sample are hybridized to the molecules in a microarray format and the signals detected after hybridization complex formation correlate to nucleic acid molecule levels in a sample. In the differential hybridization format, the differential expression of a set of genes in two biological samples is analyzed. For differential hybridization, nucleic acid molecules from both biological samples are prepared and labeled with different labeling moieties. A mixture of the two labeled nucleic acid molecules is added to a microarray. The microarray is then examined under conditions in which the emissions from the

two different labels are individually detectable. Molecules in the microarray that are hybridized to substantially equal numbers of nucleic acid molecules derived from both biological samples give a distinct combined fluorescence (Shalon *et al.* PCT publication WO95/35505). In a preferred embodiment, the labels are fluorescent markers with distinguishable excitation and emission spectra, such as Cy3 and Cy5 fluorophores.

After hybridization, the microarray is washed to remove nonhybridized nucleic acid molecules, then complex formation between the hybridizable array elements and the nucleic acid molecules is detected. Methods for detecting complex formation are well known to those skilled in the art. In a preferred embodiment, the nucleic acid molecules are labeled with a fluorescent label and measurement of levels and patterns of fluorescence indicative of complex formation is accomplished by fluorescence microscopy, preferably confocal fluorescence microscopy.

In a differential hybridization experiment, nucleic acid molecules from two or more different biological samples are labeled with two or more different fluorescent labels with different excitation and emission wavelengths. The labeled sample is excited with a specific excitation wavelength. Fluorescent signals are detected separately with different photomultipliers set to detect specific emission wavelengths. The relative abundances/expression levels of the nucleic acid molecules in two or more samples is obtained.

Typically, microarray fluorescence intensities can be normalized to take into account variations in hybridization intensities when more than one microarray is used under similar test conditions. In a preferred embodiment, individual arrayed-sample nucleic acid molecule complex hybridization intensities are normalized using the intensities derived from internal normalization controls contained on each microarray.

The labeled sample emits specific wavelengths which are detected using a plurality of photomultipliers. The relative abundances/expression levels of the arrayed nucleic acid molecules molecules can be used as hybridizable elements in a microarray. Such a microarray can be employed to identify expression profiles associated with particular toxicological responses. Then, a particular subset of these photomultipliers are set to detect specific wavelengths. The relative expression levels of the arrayed nucleic acid molecules can be identified as to which arrayed nucleic acid molecule expression is modulated in response to a particular toxicological agent. These photomultipliers set to detect specific wavelengths. The relative expression levels of the nucleic acid molecules can be employed to identify other compounds with a similar toxicological response.

Alternatively, for some treatments with known side effects, the microarray, and expression patterns derived therefrom, is employed to "fine tune" the treatment regimen. A dosage is established that minimizes expression patterns associated with undesirable side effects. This

approach may be more sensitive and rapid than waiting for the patient to show toxicological side effects before altering the course of treatment.

Generally, the method for screening a library of test compounds or molecules to identify those with a toxicological response entails selecting a plurality of arrayed nucleic acid molecule genes whose expression levels are modulated in tissues treated with known toxic compounds when compared with untreated tissues. Then a sample is treated with the test compound or molecule to induce a pattern of gene expression comprising the expression of a plurality of nucleic acid molecules. A test compound may be screened at several doses to determine which doses may be toxic and which may not.

Then, the expression levels of the arrayed nucleic acid molecules and the sample nucleic acid molecules are compared to identify those compounds that induce expression levels of the sample nucleic acid molecules that are similar to those of the arrayed nucleic acid molecules. In one preferred embodiment, expression levels are compared by contacting the arrayed nucleic acid molecules with the sample nucleic acid molecules under conditions effective to form hybridization complexes between arrayed nucleic acid molecules and sample nucleic acid molecules, and detecting the presence or absence of the hybridization complexes.

Similarity may mean that at least 1, preferably at least 5, more preferably at least 10, of the upregulated arrayed nucleic acid molecules form hybridization complexes with the sample nucleic acid molecules at least once during a time course to a greater extent than would the nucleic acid molecules of a sample not treated with the test compound. Similarity may also mean that at least 1, preferably at least 3, of the downregulated nucleic acid molecules form hybridization complexes with the nucleic acid molecules at least once during a time course to a lesser extent than would the nucleic acid molecules of a sample not treated with the test compound.

Such a similarity of expression patterns means that a toxicological response is associated with the test compound or molecule tested. Preferably, the toxic compounds belong to the class of peroxisomal proliferators (PPs), including hypolipidemic drugs, such as clofibrate, fenofibrate, clofenic acid, nafenopin, gemfibrozil, ciprofibrate, bezafibrate, halofenate, simfibrate, benzofibrate, etofibrate, WY 14,643, and the like; n-alkylcarboxylic acids, such as trichloroacetic acid, valproic acid, hexanoic acid, and the like; n-alkylcarboxylic acid precursors, such as trichloroethylene, tetrachloroethylene, and the like; azole antifungal compounds, such as bifenthrin, and the like; leukotriene D4 antagonists; herbicides; pesticides; phthalate esters, such as di-[2-ethylhexyl] phthalate, mono-[2-ethylhexyl] phthalate, and the like; and natural chemicals, such as phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol, and the like. In another embodiment, the toxic compounds are acetaminophen or one of its corresponding metabolites. In yet another embodiment, the toxic compound is a polycyclic aromatic hydrocarbon.

(PAH), including compounds such as benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, their corresponding metabolites, and the like. Of particular interest is the study of the metabolic responses of these compounds on the liver, kidney, brain, spleen, pancreas, and lung.

5     Modification of Gene Expression Using Nucleic Acids

Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the mammalian gene. Oligonucleotides designed with reference to the transcription initiation site are preferred. Similarly, inhibition can be achieved using triple helix base-pairing, which inhibits the binding of 10 polymerases, transcription factors, or regulatory molecules (Gee *et al.* In: Huber and Carr (1994) Molecular and Immunologic Approaches, Eutera Publishing, Mt. Kisco NY, pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library of nucleic acid molecules or fragments thereof may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

15       Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA followed by endonucleolytic cleavage at sites such as GUA, GUU, and GUC. Once such sites are identified, an oligonucleotide with the same sequence may be evaluated for secondary structural features which would render the 20 oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing their hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, *in vitro* or *in vivo*, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be modified to increase intracellular stability and 25 half-life by addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. The inclusion of nontraditional bases such as inosine, queosine, and wybutoxine, or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, 30 methyl-, thio- groups, renders the molecule less available as a substrate to endogenous endonucleases.

Screening Assays

The nucleic acid molecule encoding the mammalian protein may be used to screen a library of molecules for specific binding affinity. The libraries may be DNA molecules, RNA 35 molecules, PNAs, peptides, proteins such as transcription factors, enhancers, repressors, and other

ligands which regulate the activity, replication, transcription, or translation of the nucleic acid molecule in the biological system. The assay involves combining the mammalian nucleic acid molecule or a fragment thereof with the library of molecules under conditions allowing specific binding, and detecting specific binding to identify at least one molecule which specifically binds the nucleic acid molecule.

Similarly the mammalian protein or a portion thereof may be used to screen libraries of molecules in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate, or located intracellularly. Specific binding between the protein and molecule may be measured. Depending on the kind of library being screened, the assay may be used to identify DNA, RNA, or PNA molecules, agonists, antagonists, antibodies, immunoglobulins, inhibitors, peptides, proteins, drugs, or any other ligand, which specifically binds the protein. One method for high throughput screening using very small assay volumes and very small amounts of test compound is described in USPN 5,876,946, incorporated herein by reference, which screens large numbers of molecules for enzyme inhibition or receptor binding.

#### Purification of Ligand

The nucleic acid molecule or a fragment thereof may be used to purify a ligand from a sample. A method for using a mammalian nucleic acid molecule or a fragment thereof to purify a ligand would involve combining the nucleic acid molecule or a fragment thereof with a sample under conditions to allow specific binding, detecting specific binding, recovering the bound protein, and using an appropriate agent to separate the nucleic acid molecule from the purified ligand.

Similarly, the protein or a portion thereof may be used to purify a ligand from a sample. A method for using a mammalian protein or a portion thereof to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and using an appropriate chaotropic agent to separate the protein from the purified ligand.

#### Pharmacology

Pharmaceutical compositions are those substances wherein the active ingredients are contained in an effective amount to achieve a desired and intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, the therapeutically effective dose may be estimated initially either in cell culture assays or in animal models. The animal model is also used to achieve a desirable concentration range and route of administration. Such information may then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or inhibitor which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity of such agents may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 5 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it may be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indexes are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for human use.

#### Model Systems

10 Animal models may be used as bioassays where they exhibit a toxic response similar to that of humans and where exposure conditions are relevant to human exposures. Mammals are the most common models, and most toxicity studies are performed on rodents such as rats or mice because of low cost, availability, and abundant reference toxicology. Inbred rodent strains provide a convenient model for investigation of the physiological consequences of under- or over-expression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal inbred to over-express a particular gene, so that the protein is secreted in milk, may also serve as a convenient source of the protein expressed by that gene.

#### Toxicology

20 Toxicology is the study of the effects of test compounds, molecules, or toxic agents on living systems to identify adverse effects. The majority of toxicity studies are performed on rats or mice to help predict whether adverse effects of agents will occur in humans. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic, developmental, and reproductive processes, and lethality are used to generate profiles of safe or toxic responses and to assess the consequences on human health following exposure to the agent.

25 Genetic toxicology identifies and analyzes the ability of an agent to produce damage at a cellular or subcellular level. Such genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when mutated chromosomes are passed along to progeny. Toxicological studies may identify agents that increase the frequency of structural or functional abnormalities in progeny if administered to either 30 parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because of their short reproductive cycle which produces the number of organisms needed to satisfy statistical requirements.

Acute toxicity tests are based on a single administration of the agent to the subject to determine the symptomology or lethality of the agent. Three experiments are conducted: 1) an initial dose-range-finding experiment, 2) an experiment to narrow the range of effective doses,

and 3) a final experiment for establishing the dose-response curve.

Prolonged toxicity tests are based on the repeated administration of the agent. Rat and dog are commonly used in these studies to provide data from species in different taxonomic orders. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose concentrations for periods of three to four months will reveal most forms of toxicity in adult animals. Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity or the carcinogenic potential of an agent. When studies are conducted on rats, a minimum of at least one test group plus one control group are used. Animals are quarantined, examined for health, and monitored at the outset and at intervals throughout the experiment.

#### Transgenic Animal Models

Transgenic rodents which overexpress or underexpress a gene of interest may be inbred and used to model human diseases or to test compounds and molecules for therapeutic or toxicological effects. (See USPN 4,736,866; USPN 5,175,383; and USPN 5,767,337; incorporated herein by reference). In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal development or postnatally. Expression of the transgene is monitored by analysis of phenotype or tissue-specific mRNA expression, in transgenic animals before, during, and after being challenged with experimental drug therapies.

#### Embryonic Stem Cells

Embryonic stem cells (ES) isolated from rodent embryos retain the potential to form an embryo. When ES cells are placed inside a carrier embryo, they resume normal development and contribute to all tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors for knockout strains contain a disease gene candidate modified to include a marker gene which disrupts transcription and/or translation *in vivo*. The vector is introduced into ES cells by transformation methods such as electroporation, liposome delivery, microinjection, and the like which are well known in the art. The endogenous rodent gene is replaced by the disrupted disease gene through homologous recombination and integration during cell division. Then transformed ES cells are selected under conditions, identified, and preferably microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

ES cells are also used to study the differentiation of various cell types and tissues *in vitro*, such as neural cells, hematopoietic lineages, and cardiomyocytes (Bain *et al.* (1995) Dev. Biol.

168:342-357; Wiles and Keller (1991) Development 111:259-267; and Klug *et al.* (1996) J. Clin. Invest. 98:216-224). Recent developments demonstrate that ES cells derived from human blastocysts may also be manipulated *in vitro* to differentiate into eight separate cell lineages, including endoderm, mesoderm, and ectodermal cell types (Thomson (1998) Science 282:1145-5 1147).

#### Knockout Analysis

In gene knockout analysis, a region of a human disease gene candidate is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (*neo*; Capecchi (1989) Science 244:1288-1292). The inserted coding sequence disrupts transcription 10 and translation of the targeted gene and prevents biochemical synthesis of the disease candidate protein. The modified gene is transformed into cultured embryonic stem cells (described above), the transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines.

#### Knockin Analysis

15 Totipotent ES cells, present in the early stages of embryonic development, can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human sequence integrates into the animal cell genome by recombination. Totipotent ES cells which contain the integrated human gene are handled as described above. Inbred animals are 20 studied and treated to obtain information on the analogous human condition. These methods have been used to model several human diseases. (See, e.g., Lee *et al.* (1998) Proc. Natl. Acad. Sci. 95:11371-11376; Baudoin *et al.* (1998) Genes Dev. 12:1202-1216; and Zhuang *et al.* (1998) Mol. Cell Biol. 18:3340-3349).

#### Non-Human Primate Model

25 The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in evaluating the effects of test compounds or molecules on non-human primates as they can be related to human health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures under similar conditions. Cynomolgus and 30 rhesus monkeys (*Macaca fascicularis* and *Macaca mulatta*, respectively) and common marmosets (*Callithrix jacchus*) are the most common non-human primates (NHPs) used in these investigations. Since great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, 35 NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be

classified as "extensive metabolizers" and "poor metabolizers" of these agents.

In additional embodiments, the nucleic acid molecules which encode the mammalian protein may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleic acid molecules that are currently known, 5 including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

### Examples

It is understood that this invention is not limited to the particular methodology, protocols, 10 and reagents described, as these may vary. It is also understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims. The examples below are provided to best describe the subject invention and its representative constituents.

#### I cDNA Library Construction

15 The RALINOT01 cDNA library was constructed from liver tissue removed from a pool of fifty 10- to 11-week-old Sprague-Dawley female rats (Pharmacon, Waverly PA). The animals were housed in standard laboratory caging and fed PMI-certified Rodent Diet #5002. The animals appeared to be in good health at the time tissue was harvested. The animals were anesthetized by CO<sub>2</sub> inhalation, and then cardiotocesis was performed.

20 Frozen tissue was homogenized and lysed in TRIZOL reagent (1 g tissue/10 ml TRIZOL; Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate, using a POLYTRON homogenizer (PT-3000; Brinkmann Instruments, Westbury NY). After a brief incubation on ice, chloroform (1:5 v/v) was mixed with the reagent, and then centrifuged at 1,000 rpm. The upper aqueous layer was removed to a fresh tube, and the RNA precipitated with 25 isopropanol, resuspended in DEPC-treated water, and treated with DNase I for 25 min at 37°C. The RNA was re-extracted once with phenol-chloroform, pH 4.7, and precipitated using 0.3 M sodium acetate and 2.5 volumes ethanol. The mRNA was then isolated using an OLIGOTEX kit (QIAGEN, Chatsworth CA) and used to construct the cDNA library.

30 The mRNA was handled according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL-4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into the pINCY1 plasmid vector (Incyte Pharmaceuticals). The plasmid pINCY1 was subsequently transformed into DH5α or DH10B competent cells (Life Technologies).

35 The RAKINOT01 library was constructed using mRNA isolated from kidney tissue removed from a pool of fifty, 7- to 8-week-old male Sprague-Dawley rats, as described above.

The RAKINOT02 library was constructed using mRNA isolated from kidney tissue removed from a pool of fifty, 10- to 11-week-old female Sprague-Dawley rats, as described above.

## II cDNA Library Normalization

In some cases, cDNA libraries were normalized in a single round according to the procedure of Soares *et al.* (1994, Proc. Natl. Acad. Sci. 91:9228-9232) with the following modifications. The primer to template ratio in the primer extension reaction was increased from 2:1 to 10:1. Reduction of each dNTP concentration in the reaction to 150 $\mu$ M allowed the generation of longer (400-1000 nucleotide (nt)) primer extension products. The reannealing hybridization was extended from 13 to 19 hours. The single stranded DNA circles of the normalized library were purified by hydroxyapatite chromatography, converted to partially double-stranded by random priming, and electroporated into DH10B competent bacteria (Life Technologies).

The Soares normalization procedure is designed to reduce the initial variation in individual cDNA frequencies and to achieve abundances within one order of magnitude while maintaining the overall sequence complexity of the library. In the normalization process, the prevalence of high-abundance cDNA clones decreases significantly, clones with mid-level abundance are relatively unaffected, and clones for rare transcripts are increased in abundance. In the modified Soares normalization procedure, significantly longer hybridization times are used to increase gene discovery rates by biasing the normalized libraries toward low-abundance cDNAs that are well represented in a standard transcript image.

The RALINON03, RALINON04, and RALINON07 normalized rat liver cDNA libraries were constructed with  $2.0 \times 10^6$ ,  $4.6 \times 10^5$ , and  $2.0 \times 10^6$  independent clones from the RALINOT01 cDNA library, respectively. The RALINOT01 cDNA library was normalized in one round using conditions adapted from Soares (*supra*) except that a significantly longer (48-hour) reannealing hybridization was used.

## III cDNA Library Prehybridization

The RALINOH01 cDNA library was constructed with clones from the RALINOT01 cDNA library. After preparation of the RALINOT01 cDNA library, 9,984 clones were spotted onto a nylon filter, lysed, and the plasmid DNA was bound to the filter. The filter was incubated with pre-warmed hybridization buffer and then hybridized at 42°C for 14-16 hours in 0.75 M NaCl, 0.1 M Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>, 0.15 M tris-HCl (pH 7.5), 5x Denhardt's Solution, 2% SDS, 100  $\mu$ g/ml sheared salmon sperm DNA, 50% formamide, and [<sup>32</sup>P]-labeled oligonucleotide molecules made from reverse transcribed rat liver mRNA from an untreated animal. The filter was rinsed with 2 x SSC (saline sodium citrate) at ambient temperature for 5 minutes followed by washing for 30 minutes at 68°C with pre-warmed washing solution (2 x SSC, 1% SDS). The wash was

repeated with fresh wash solution for an additional 30 minutes at 68°C. Filters were then washed twice with pre-wash solution (0.6 x SSC, 1% SDS) for 30 minutes at 68°C. Some 4,224 clones had very strong hybridization signals and about 20% of the clones had no signals and two groups were isolated and sequenced.

5      **IV      Isolation and Sequencing of cDNA Clones**

DNA was isolated using the following protocol. Single bacterial colonies were transferred into individual wells of 384-well plates (Genetix Ltd, Christchurch, United Kingdom) using sterile toothpicks. The wells contained 1 ml of sterile Terrific Broth (Life Technologies) with 25 mg/l carbenicillin and 0.4% glycerol (v/v). The plates were covered and placed in an incubator (Thermodyne, Newtown Square PA) at 37°C for 8-10 hours. Plasmid DNA was released from the cells and amplified using direct link PCR (Rao, V.B. (1994) Anal. Biochem. 216:1-14) as follows. The direct link PCR solution included 30 ml of NUCLEIX PLUS PCR nucleotide mix (Amersham Pharmacia Biotech, Piscataway NJ) and 300 µl of Taq DNA polymerase (Amersham Pharmacia Biotech). Five microlitres of the PCR solution were added to each of the 384 wells using the 10      MICROLAB 2200 system (Hamilton, Reno NV); plates were centrifuged at 1000 rpm for 20 seconds and refrigerated at 4°C. A 24 pin tool (V&P Scientific Inc, San Diego CA) was used 15      to transfer bacterial cells from the original plate into the plate containing the PCR solution where 0.1% Tween 20 caused the cells to undergo lysis and release the plasmid DNA. After lysis, the plates were centrifuged up to 500 rpm, covered with a cycle sealer, and cycled using a 384-well DNA ENGINE thermal cycler (MJ Research, Watertown MA) using the program dPCR30 with the following parameters: Step 1) 95°C, 1 minute; Step 2) 94°C, 30 seconds; Step 3) 55°C, 20      30 seconds; Step 4) 72°C, 2 minutes; Step 5) steps 2, 3, and 4 repeated 29 times; Step 6) 72°C, 10 minutes; and Step 7) storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICO 25      GREEN quantitation reagent (0.25% (v/v), Molecular Probes, Eugene OR) dissolved in 1x TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the quantitation reagent. The plate was scanned in a Fluoroscan II (Labsystems Cy, Helsinki, Finland) to measure the fluorescence of the sample and to 30      quantitate the concentration of DNA. Typical concentrations of each DNA sample were in the range of 100 to 500 ng/ml.

The cDNAs were prepared for sequencing using either a HYDRA microdispenser (Robbins Scientific, Sunnyvale CA) or MICROLAB 2200 system (Hamilton) in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were sequenced using the method of Sanger, F. and A.R. Coulson (J. Mol. Biol. (1975) 94:441-448) and the ABI 377 sequencing 35      systems (PE Biosystems). Most of the isolates were sequenced according to standard ABI

protocols using ABI kits (PE Biosystems). The solution volumes were used at 0.25x - 1.0x concentrations. Typically, 500 to 700 base pairs were sequenced in 3.5 to 4 hours. In the alternative, cDNAs may have been sequenced using solutions and dyes from Amersham Pharmacia Biotech.

5      **V       Rat Liver and Kidney Gene Selection**

As a first step, originator molecules from high throughput sequencing experiments were derived from clone inserts from RALINOT01, RAKINOT01, RAKINOT02, RALINOH01, RALINON03, RALINON04 and RALINON07. cDNA library clones were obtained. There were 18,140 rat liver molecules and 5,779 rat kidney molecules.

10     Additionally, 1,500 rat molecules derived from clone inserts of any of 113 rat cDNA libraries were selected based on their homology to genes coding for polypeptides implicated in toxicological responses including peroxisome-associated genes, lysosome-associated genes, apoptosis-associated genes, P450 cytochromes, detoxification genes such as sulfotransferases, glutathione S-transferase, and cysteine proteases, and the like.

15     Then, all the remaining molecules derived from all of the rat cDNA library clones were clustered based on the originator molecules described above. The clustering process involved identifying overlapping molecules that have a match quality indicated by a product score of 50 using BLAST.

6581 master clusters were identified.

20     After forming the clone clusters, a consensus sequence was generated based on the assembly of the clone molecules using Phrap (Phil Green, University of Washington). The assembled molecules were then annotated by first screening the assembled molecules against GenBank using BLASTn and then by screening the assembled molecules against GenPept using FASTX. About two thirds of the assembled molecules were annotated, about one third of the 25 assembled molecules were not annotated.

**VI       Substrate and Array Element/Probe Preparation**

Clones nominated in the process described in Example V were used to generate array elements. Each array element was amplified from bacterial cells. PCR amplification used primers complementary to the vector sequences flanking the cDNA insert. Array elements were amplified 30 in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements were then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements were immobilized on polymer-coated glass slides. Glass microscope slides (Corning, Corning NY) cleaned by ultrasound in 0.1% SDS and acetone, with 35 extensive distilled water washes between and after treatments. Glass slides were etched in 4%

hydrofluoric acid (VWR, West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma-Aldrich, St. Louis MO) in 95% ethanol. Coated slides were cured in a 110°C oven.

5 Array elements were applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522 and incorporated herein by reference. In brief, 1 µl of the array element DNA, at an average concentration of 0.5 µg/ml in 3 x SSC, was loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposited about 5 nl of the array element sample per slide. A total of 7404 array elements representing rat liver and kidney genes and a variety of control elements, including 14 synthetic control molecules, human genomic 10 DNA, and yeast genomic DNA, were arrayed in four identical quadrants within a 1.8 cm<sup>2</sup> area of the glass substrate.

15 Microarrays were UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays were washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites were blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

## VII. Target Preparation

Male Sprague-Dawley rats (6-8 wk old) were dosed intraperitoneally with clofibrate (CLO; Acros, Geel, Belgium) at 250 mg/kg body weight (bw), acetaminophen (APAP; Acros) at 20 1000 mg/kg bw, benzo(a)pyrene (B(a)P; Acros) at 10 mg/kg bw, or dimethylsulfoxide vehicle (DMSO; Acros) at less than 2 ml/kg bw and the animals were later euthanized by CO<sub>2</sub> inhalation. Animals were monitored daily for physical condition and body weight. Three animals per group were sacrificed approximately 12 hours, 1 day (d), 3d, 7d, 14d, and 28d following the single dose. Prior to sacrifice a blood sample from each animal was drawn and assayed for serum alanine 25 transferase (ALT) and aspartate aminotransferase (AST) levels using a diagnostic kit (Sigma-Aldrich). Observed gross pathology and liver weights were recorded at time of necropsy. Liver, kidney, brain, spleen and pancreas from each rat were harvested, flash frozen in liquid nitrogen, and stored at -80°C.

For each probe preparation, frozen liver was homogenized and lysed in TRIZOL reagent 30 (Life Technologies, Gaithersburg MD) following the modifications for liver RNA isolation. Messenger RNA was isolated using an OLIGOTEX kit (QIAGEN) and labeled with either Cy3- or Cy5-labeled primers (Operon Technologies, Alameda CA) using the GEMBRIGHT labeling kit (Incyte Pharmaceuticals). Messenger RNA isolated from tissues of rats treated with clofibrate, acetaminophen, or benzo(a)pyrene was labeled with Cy5 and mRNA isolated from tissues of rats 35 treated with DMSO was labeled with Cy3. Quantitative and differential expression pattern control

cDNAs were added to each labeling reaction. Labeled cDNA was treated with 0.5 M sodium bicarbonate (pH 9.2) for 20 min at 85 °C to degrade the RNA and purified using two successive CHROMA SPIN 30 gel filtration spin columns (Clontech, Palo Alto CA). Cy3-labeled control sample and Cy5-labeled experimental sample were combined and precipitated in glycogen, sodium acetate, and ethanol.

Probes are also prepared from tissue needle biopsy samples. Samples are used to identify changes within the tissue following exposure to, for example, a toxic compound, a potential toxic compound, a compound with unknown metabolic responses, or a pharmacological compound.

## VIII Hybridization

10 Hybridizations were carried out using the methods described by Shalon (*supra*).

IX Detection

The microscope used to detect the reporter-labeled hybridization complexes was equipped with an Innova 70 mixed gas 10 W laser (Coherent Lasers, Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3, and 632 nm for excitation of Cy5. The excitation laser light was focused on the array using a 20x microscope objective (Nikon, Melville NY). The slide containing the array was placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example was scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excited the two fluorophores sequentially. Emitted light was split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics, San Jose CA) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes were used to filter the signals. The emission maxima of the fluorophores used were 565 nm for Cy3 and 650 nm for Cy5. Each array was typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus was capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans was typically calibrated using the signal intensity generated by a cDNA control species added to the probe mix at a known concentration. A specific location on the array contained a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two probes from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration was done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

35 The output of the photomultiplier tube was digitized using a 12-bit RTI-835H analog-to-

digital (A/D) conversion board (Analog Devices, Norwood MA) installed in an IBM-compatible PC computer. The digitized data were displayed as an image where the signal intensity was mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data was also analyzed quantitatively. Where two different 5 fluorophores were excited and measured simultaneously, the data were first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid was superimposed over the fluorescence signal image such that the signal from each spot was centered in each element of the grid. The fluorescence signal within each element 10 was then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis was the GEMTOOLS gene expression analysis program (Incyte Pharmaceuticals). In one analysis, where two different samples were prepared from identically treated cell cultures, expression patterns of those cDNAs which changed between 1.6- and 1.7-fold were within the 95% confidence limits of a Poisson normal distribution profile 15 (T. Theriault, pers. communication).

#### X Results

The expression patterns of eight cytochrome P450 isozymes known to be induced in a toxicological response were monitored during the 28 day time course. The results using clofibrate, acetaminophen, and benzo(a)pyrene are shown in Table 1, Table 2, and Table 3, respectively. 20 Each of the known genes was upregulated greater than 2 fold at least once during the time course.

TABLE 1 Gene expression patterns (x fold increase) of known genes in clofibrate-treated rat liver

Gene	12 hours	1 day	3 days	7 days	28 days
P450 LA-omega	15	26	2.0	2.1	3.0
P450 4A	6.5	16.5	2.1	3.0	3.5
P450 3A	0.14	1.6	0.63	0.50	0.45

30 TABLE 2 Gene expression patterns (x fold increase) of known genes in acetaminophen-treated rat liver

Gene	12 hours	24 hours	3 days	7 days	14 days	28 days
P450A	1	4.4	2.2	2.0	4.6	4.8
P450F	0.50	0.23	2.0	1.8	2.2	2.2
P450 14DM	0.45	0.32	2.2	1.6	1.8	0.56

TABLE 3 Gene expression patterns (x fold increase) of known genes in benzo(a)pyrene-treated rat liver

Gene	12 hours	1 day	3 days	7 days	14 days	28 days
P450 LA-omega	1.2	2.3	2.4	1.4	6.8	1.2
P450 MCA-inducible	8.2	11.8	4.4	2.2	2.4	1.2
P450 ISF/B-NF	9.6	7.4	6.2	2.4	2.4	1.2

We have discovered novel nucleotide molecules that are up-regulated or down-regulated at least 2-fold at least once during the time course. These molecules are SEQ ID NOs:1-117 provided in the Sequence Listing. These polynucleotide molecules can be used for screening test compounds or molecules for a toxicologic effect.

Table 4 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with clofibrate (CLO) and Table 5 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with clofibrate.

TABLE 4 Gene expression patterns (x fold increase) of CLO-upregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	28 days
35	11.6	14.4	2.4	3.0	3.2
36	11.6	18.7	3.0	3.3	3.8
31	1.2	2.8	1.0	2.3	4.8
57	0.9	1.9	0.9	1.5	4.5
67	4.3	1.1	1.6	1.7	5.7
81	5.1	1.2	1.7	1.8	6.0
94	4.8	1.4	2.0	1.5	2.4
33	5.1	1.3	1.9	1.8	5.5

TABLE 5 Gene expression patterns ( $\times$  fold increase) of CLO-downregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	28 days
5	102	0.15	1.4	1.0	0.77
	103	0.13	1.2	0.83	0.63
	52	0.13	0.56	0.37	1.0
	43	0.13	1.1	0.91	0.71
	53	0.11	0.67	0.36	1.0
10	54	0.14	0.63	0.59	1.1
	55	0.16	0.67	0.71	1.2
	63	0.33	0.14	1.1	0.83
	105	0.14	1.2	1.0	0.77
	68	0.16	0.67	0.53	1.1
15	71	0.43	0.18	0.40	0.34
	74	0.06	0.71	0.42	1.1
	115	0.22	1.5	0.77	1.7
	85	0.19	0.45	1.0	1.3
	90	0.12	0.48	1.2	1.0
20	95	0.14	0.91	0.56	1.5
					1.4

20

Table 6 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with acetaminophen (APAP) and Table 7 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with acetaminophen.

25

30

TABLE 6 Gene expression patterns (x fold increase) of APAP-upregulated nucleic acid molecules

SEQ ID NO:	12 hours	24 hours	3 days	7 days	14 days	28 days
5	35	3.1	6.6	2.9	3.3	4.9
	36	4.7	10.1	4.0	4.2	6.9
	78	0.9	4.4	1.2	1.5	1.1
	81	2.9	5.1	1.4	1.8	2.3
	82	1.2	4.2	1.3	1.0	1.7
	39	2.4	9.0	2.6	1.7	2.2
10	94	1.2	4.9	1.2	1.1	2.0
	33	4.3	5.9	1.5	1.7	2.9
	98	1.3	6.1	1.5	1.9	1.8
						2.1

TABLE 7 Gene expression patterns (x fold increase) of APAP-downregulated nucleic acid  
15 molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
20	49	0.59	0.15	1.2	1.0	0.83
	50	0.83	0.37	0.43	0.37	0.22
	52	0.63	0.08	1.0	0.71	0.83
	53	0.25	0.07	1.1	0.71	0.83
	54	0.43	0.19	0.04	0.71	0.29
	55	0.35	0.22	0.07	0.77	0.31
	56	0.38	0.21	0.5	0.32	1.1
	59	0.18	0.77	2.5	1.4	1.2
	61	0.15	0.53	0.91	0.71	0.71
	63	0.13	0.05	0.23	0.77	0.43
	74	0.19	0.09	1.1	1.0	1.4
	87	0.42	0.10	0.53	0.63	0.63
30	90	0.16	0.29	1.2	0.77	0.83
	95	0.22	0.20	2.7	1.7	1.6
						1.0

Table 8 shows the gene expression pattern of selected molecules that were upregulated at least 2-fold at least once during the time course following treatment with benzo(a)pyrene (B(a)P) and Table 9 shows the gene expression pattern of selected molecules that were downregulated at least 2-fold at least once during the time course following treatment with benzo(a)pyrene.

5

TABLE 8 Gene expression patterns of B(a)P-upregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
10	3	3.4	19	0.7	0.5	1.99
	9	1.6			1.1	3
	10	2.8			2.1	2.9
	13	2.9		3.1	2.3	3.3
	19	2.7	3.5	3	1.9	1.7
	26	1.1	4.7	1.5	1.3	5
	31	2.3	3.8	1.6	2	1.7
	33	2.1	4.1	3.2	2	1.7
	35	1.2	3	5.1	1.4	5
	37	3.4	0.5	0.6	0.7	0.9
15	39	1.5	3.5	1.8	1.5	3.5
	42	9.1	9.1	5.2	2.4	2.1
20						1

TABLE 9 Gene expression patterns of B(a)P-downregulated nucleic acid molecules

SEQ ID NO:	12 hours	1 day	3 days	7 days	14 days	28 days
25	11	0.4		0.4	0.3	0.53
	25	0.3		0.5	0.7	0.42
	27	1	0.1	1	1.1	0.09
	28	0.3	0.3	1.2	1.2	0.77
	45	1.2	0.2	0.4	0.6	0.77

30

## CLAIMS

What is claimed is:

1. A method for detecting or diagnosing the effect of a test compound or molecule associated with increased or decreased levels of a nucleic acid molecule in a mammalian subject comprising:
  - 5 a) treating a mammalian subject with a toxic compound or molecule;
  - b) obtaining a sample containing nucleic acids from the mammalian subject treated with the toxic compound or molecule;
  - c) contacting the sample with a microarray comprising a plurality of nucleic acid molecules comprising SEQ ID NOs: 1-117, or a fragment thereof, under conditions for the formation of one or more hybridization complexes;
  - d) detecting the hybridization complexes, wherein the presence, absence or change in amount of the hybridization complex, as compared with the hybridization complexes formed from nucleic acid molecules from an untreated mammalian subject, is indicative of a metabolic response to the toxic compound or molecule;
- 15 e) measuring the level of nucleic acid molecules in a sample from a mammalian subject treated with a test compound or molecule using the method of steps (c) and (d); and f) comparing the level detected in step (e) to a level of nucleic acid molecules present in normal or untreated biological sample in which an increase or decrease in the level of nucleic acid molecule as compared to normal levels indicates a toxicological response.
- 20 2. The method of claim 1 wherein the toxic compound or molecule is selected from hypolipidemic drugs, n-alkylcarboxylic acids, n-alkylcarboxylic acid precursors, azole antifungal compounds, leukotriene D4 antagonists, herbicides, pesticides, phthalate esters, phenyl acetate, dehydroepiandrosterone sulfate, oleic acid, methanol and their corresponding metabolites, acetaminophen and its corresponding metabolites, benzo(a)pyrene, 3-methylcholanthrene, benz(a)anthracene, 7,12-dimethylbenz(a)anthracene, and their corresponding metabolites.
- 25 3. The method of claim 1 wherein the sample is a tissue selected from the group consisting of liver, kidney, brain, spleen, pancreas, and lung.
4. The method of claim 1 wherein the test compound which elicits the metabolic response is a compound with previously unknown metabolic response.
- 30 5. The method of claim 1 wherein the test compound or molecule which elicits the metabolic response induces at least a 2-fold change in the amount of the hybridization complexes formed with at least one of the nucleic acid molecules of the sample.
6. An isolated and purified nucleic acid molecule selected from SEQ ID NOs: 1-117, or a fragment thereof.
- 35 7. A method of using the nucleic acid molecule of claim 6 to screen a library of molecules or

compounds to identify at least one molecule or compound which specifically binds the nucleic acid molecule, the method comprising:

a) combining the nucleic acid molecule of claim 6 with a library of molecules or compounds under conditions to allow specific binding; and

5 b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the nucleic acid molecule.

8. The method of claim 7 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides, and proteins.

## SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.  
CUNNINGHAM, Mary Jane  
ZWEIGER, Gary B.  
PANZER, Scott R.  
SEILHAMER, Jeffrey J.

<120> TOXICOLOGICAL RESPONSE MARKERS

<130> PA-0010 PCT

<140> To Be Assigned  
<141> Herewith

<150> 09/141,825; 09/172,711; 09/172,108  
<151> 1998-08-28; 1998-10-13; 1998-10-13

<160> 117

<170> PERL Program

<210> 1  
<211> 259  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700416862F7

<400> 1  
gtggcggcga tttctgcgtc gagcatttg agtttctcg ctgctgaacg gtagactaa 60  
acggcggctg acatggtgg a ggggtacag a a g c a t t c t g t g c a c a c a c t a g t g t t c a g g 120  
tcattgaaga ggacctatga catgtttgtg gctgataatg gaaaacctgt gcctttggat 180  
gaagagagtc acaaggcgaa aatggcaatc a a g c t t c g t a a t g a g t a t g g c c c t g t g c t g 240  
catatgccta cttcaaaag 259

<210> 2  
<211> 295  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 212, 227, 229, 232, 240, 243, 245, 250, 257, 267, 269, 273, 277  
<222> 288, 290  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700502957F6

<400> 2  
gcctcttcca ccatccggcc tagtcactgc agggggccatg cctacacctt ccactcaact 60

tgttacctct gcggctccag gcagggctta gtccaacctg cccagacacg gttcacctt 120  
 ttatgcccaa gcttcgggg tgctgaggta ggggctgcct tcctgcaccc ccaaggagca 180  
 gacactcaag aatggagtca gcttaggaacc cngggagctg cctcatnang cnittgatan 240  
 cangnacacn tttgcanctg cagaccntnt tcngaanaac nttgccangn tcaac 295

<210> 3  
 <211> 273  
 <212> DNA  
 <213> Rattus norvegicus

<220>  
 <221> unsure  
 <222> 56, 220, 235, 237, 239, 249, 256  
 <223> a or g or c or t, unknown, or other

<220>  
 <221> misc\_feature  
 <223> Incyte template ID No: 700504502F6

<400> 3  
 tgaccctgct tgctcaggt gactggtcaa gtgcgagcta gcttagttag tgtggngtat 60  
 aaggcgccat cattcctcca gtaagcctcc atcccaaagc aactgaggct gtggcagtga 120  
 tgccagcaac ctgtgtcacc caaaattatac cagccctcca cgggcactgc ctaggacctg 180  
 gggagggaaag ggactttgca tcacatagcc tcaggttcgn gtttggctct ggtangngnt 240  
 gcctgaaant ggtggnttcc agctggtgta cg 273

<210> 4  
 <211> 264  
 <212> DNA  
 <213> Rattus norvegicus

<220>  
 <221> misc\_feature  
 <223> Incyte template ID No: 700505170F6

<400> 4  
 ggctggctcg cgatggcccg ctacctcgcc tcgctggaac catgtgtggg tccggcactt 60  
 gagactggaa tcctgaaagg ggtgaacctt cagcggaaac ttgceggaaa ttttactccg 120  
 tccggacagc cacggcggga ggaggcagtg aatgcttgc gctggggcac aggccggcag 180  
 acccagattt tggtgggatg tgccggacagg accgtgaggc actttaatgc ggaggagggt 240  
 acattccaga ccagagatac tgcc 264

<210> 5  
 <211> 268  
 <212> DNA  
 <213> Rattus norvegicus

<220>  
 <221> unsure  
 <222> 231  
 <223> a or g or c or t, unknown, or other

<220>  
 <221> misc\_feature

<223> Incyte template ID No: 700511170F6

<400> 5

ggcaacaaga cgctgtgatt ggaagcaatg acgaagtccct cacactccgg ggagtgggta 60  
tgtgctgcta catcatgtga tgggcagcct ggagggatg cagggcgccct ggagctatgt 120  
ccagggtggc atgggtgccc tctcagatgc catttgcaga tcggctactg cacatggagc 180  
aagtatottc acagagaaga ctgtggctaa ggtgcaagtg aacagcgaag ncgtgtccaa 240  
ggggtcgtgc tcagggccgg cgaggagt 268

<210> 6

<211> 284

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 28, 37, 44, 67, 71, 88, 102, 172, 238

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700607827F7

<400> 6

gtgtacctac ctgctgagga aggaaggngt ggatgnacc cagnaacctg atgtccagca 60  
caagggnAAC nggcgtggct ttctaatnta caagcctggg tnactacagc ttgcagctac 120  
ctaacccatg caggaggcga accctctgag cccagttgct attgtgacca tnaagatgtc 180  
ttgccacaca gcttcacccc agtctgggtt taatggaaag ttacctaacg attaccnca 240  
gaagacacat gagacgcttg cttcaaagct ctcggatgca gccA 284

<210> 7

<211> 243

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte template ID No: 700500814F6

<400> 7

actttttctta atgtcttatg gccacttctt atgagaatgg ggagctgctc tgccctgaggg 60  
tgttgagagg aagcggcaga gcaggccat catcccaacc ctggccttg gccctcccc 120  
ctagctctgc agcatttctt cagatcctct ttcctgagag tcaaggagac taaacaccaa 180  
taaaccagac acaacccctcg tggccccaAA ggagaaaccg attagagggt tctctgttag 240  
atg 243

<210> 8

<211> 259

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 235

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700526819F6

<400> 8

ggcgcaggcg gaaggggcct gtcaccgtcc gctgcgacgt cgccggctgga gttgaacctg 60  
gtgccggctg ctttgcgtg tgagtgcgtg gcccggaaa aacgagaacc ggacgagtgg 120  
cgctctggaga agtatgtggg tccttggaa acatgcgtca gccctgaaag tccaagcaag 180  
taaacccgcct cggaagtgtat cagtgagtaa tcccgcaaag tgactttctg aaggnatgct 240  
gaggctgaga agtgactct 259

<210> 9

<211> 255

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 2, 4, 53, 55, 86, 88, 92, 106, 133, 211, 214

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700528176F6

<400> 9

cnaangggcct caggaaatca agaggagcca gcctgatccc tggcctctgg agncntaaaa 60  
caagtgttgtt tttgcaggta gcctangntg gntgtcgatg gagctncagc ctgcattggca 120  
tttaggcagga agncactctg gatgattgtg cacatgagaa cctagtcagg gagggagggt 180  
ttaaggagag gcttagaata caagtgagaa ncancgagaa agggaccaag tcctcagaat 240  
agaagctatc tgcct 255

<210> 10

<211> 269

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 30-31, 140

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700528082F6

<400> 10

ataaaaagtga aaactggca agggcaaggn ngctggcgt gaaccgctta ctagataatg 60  
gtctctaaaa attggctctg aaaaccctgt ttgtgttattc gttttatgag tgcttaaaaa 120  
tggtgtgacc agggcatggn cactgtcatt ggaacagcaa catgtttgct ggcacattgg 180  
aatggggaaa tgtgaagaaa gctggcatca ggccctgcggc acccatttct ttgatgaaag 240  
tggtgtgtca aaccccaact aatcatttt 269

<210> 11  
<211> 254  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 36, 47, 67, 80, 82-83, 92, 111  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700535328F6

<400> 11  
caaggcgttc tgctgcgaga acgacattga catcgngcgc gtggtanacg tgcggaggct 60  
ggcggnatc gtggcgccn annacgagag gngcgcgccg ttagacttgc nttgcattct 120  
catttcgaac cctaattgaa acacatggaa ggaccctgcc ttggagaagc tcagtttgtt 180  
ctgcgaggag agccgcagct tcaacgactg ggtcccagca tcacccttcc gagtgacagc 240  
ctgcaggac ctgg 254

<210> 12  
<211> 244  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 50, 102, 169, 171, 204, 210, 234, 242  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700368973F6

<400> 12  
cacaatccca aactggaaaa actaaaaaag gaatccctgct gtgaaaggtn tatattactc 60  
tagatttttc ttactgtaaa tattgttaaga ttgttaact gncaatattt tattaaccaa 120  
caaatgttaa tctatgtgaa atcagactta tttaaaggc tgctattang ngtgtggccc 180  
tttgctgaca gattaagtat attntgagtn agataactta ttaaggatgg aacnttaaag 240  
gnnc 244

<210> 13  
<211> 237  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 11, 20, 132, 136, 149, 151, 158, 171, 173, 184, 187, 190, 193  
<222> 204-205, 208, 212-213, 215, 221, 227, 230-231, 236  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature

<223> Incyte template I No: 700368974F6

<400> 13

gtccttagct ngtggcggtt ggggtgcagt ctgctactat ctgaacgaat ttaatgtggg 60  
agcatgcctt atacaacacg ggaaattaat gtgtgatcta atgcgtgatc tatgacttat 120  
tacaatacag anttangtgtt gAACCTGCNT ncaaaacngg tcagaatttt ngnaatggcc 180  
ggantgnaon ggntgnttat taanntgnaa gnngntggga naggccnggn ntgcgnt 237

<210> 14

<211> 235

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 38, 49, 53, 57, 125, 132, 136, 138, 150-151, 163, 165-166

<222> 179-180, 184, 185, 189, 192, 199, 219-221, 223

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700369461F6

<400> 14

ctttaaccggg tgggctgctg taagaatcgg tggcaggnc ctctctgcng ggnntaatt 60  
gctctgaaac gctacttagga cccgaatact aaggccacat ctctacngtc taagagggga 120  
aatangatag ctttgtncc acatgtggcn nagtgggtt gcngnntatin gcttaacann 180  
tacnanttnc antgattant gtgggtgtaa gatggctnn ntnaaaaactg ccgcc 235

<210> 15

<211> 205

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 17, 19, 21, 56, 69, 75, 81, 103, 109, 111, 122, 127, 130, 137

<222> 148, 151, 153, 166, 174, 177, 190, 193, 195

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700782579F6

<400> 15

gggtcattta caacctntna nacaagggga cgccccaga catgccagtg ttcacngaac 60  
aagatgcang tccancagga ncagatagac tcagtcctatg gantggctna ncaaccagcc 120  
cngggcngcn caactgntgg acangtangg nantttctgc gacaantggg gagnttnttg 180  
cgaagctcon tgncngaaag aatgg 205

<210> 16

<211> 236

<212> DNA

<213> Rattus norvegicus

WO 00/12760

<220>  
<221> unsure  
<222> 11-12, 33-34, 46, 76, 85-86, 103, 123, 127, 144, 152, 157, 162  
<222> 169, 174, 183-184, 186, 191-194, 217, 224, 230  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700782665F6

<400> 16  
ggctccatg nngaccgggc gccttgggt tgnngagacg ctgcangccc ttaacgcccc 60  
ctttagggc cttgnaacc ccgggnntgg ttaaggaaaa cgnatgcccc acaccccg 120  
gtncganact ttttggcacc gcnnaaactc gnttacngtg gnttnang tagngggtat 180  
gttncncgag nnntttagg ccggcttgtg ctgcggatt gggntgaaan tgtctg 236

<210> 17  
<211> 267  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700480927F6

<400> 17  
ctggaccaac atcacaagaa atgaataaaag cagatttctc tggtgagttc tgcagtaaac 60  
cacctaaaag ccaatgtcaa gtcagccgca gacttactta gcctgcctag cactgttagag 120  
ggacttcaga agagtgtcgc ttccattggc aatacgttga acagtgtcag cttgctgta 180  
gaggcaatac agaagaccgt ggatgaacac aaggcacctt ggagttactg cagggcagtg 240  
tggagaccaa tggagcaac caaatca 267

<210> 18  
<211> 271  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 2, 92, 162, 181, 247-248, 253, 256, 262, 269  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700481732F6

<400> 18  
gncttattta tgtatgaaaa tgcagaaatc tgtacattcc tcaagccagt cctgtcgagc 60  
caggctgtc ccatccttgt acctcaaccc antcccacct ggcctgaaca ccccatgaga 120  
cagagctgt ctctggctg gggccccccag gcctggctg gncaggcaga ccctaccccg 180  
nagtccactg gtcctcgtct ccgaggctct cctggctac aaagggggac cacacacacc 240  
cagaatnntt tantgnattg gngggcccnng g 271

<210> 19

<211> 283  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 37, 131, 199, 205-206, 208, 232-233, 239, 244-246, 251, 253-254  
<222> 261-262, 266, 269, 274-275, 280  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700483143F6

<400> 19  
ctaaaattaa gatagaagtg aatgagacag atatctngta agacactgtt ttttcttg 60  
tgatcagatc tagtgtggtg ggatgataga agttgaactt gctttattgc tatgggttaa 120  
aatatttgt ntcattaaaa tggcctattt aaatgcttt ctgttccat aataaaataa 180  
cctgatgaaa aagtaaaaana aaaannanaa aaaaaaaggg gcggcccccg cnnaggcnt 240  
tttnnncccg ngnnttattt nngccnggnc cctrnngggn cca 283

<210> 20  
<211> 256  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700484538F6

<400> 20  
ctggcctcag cttcctaagt tctggagtc ggacagggtt ttgccacaca catggccac 60  
cggggaccta gaacctacag tgaaccgtca cccaggtct gtggatgttc tgcatcctga 120  
ggtagacagc ctctaataatc ctgttaggga cctaggacca gagctgggtt gcccaggcat 180  
gtccccacat gtcgcacatcgg ccacagggat atcggttgaa gtgcatttgg aagtgtgctg 240  
ggacgccagc cagctt 256

<210> 21  
<211> 272  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 64, 68, 124, 140, 145, 213, 260, 267, 269  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700480077F6

<400> 21  
ctgacacctac ccatttttta aggaccgtt tagcacgg accactgcaa aggccggctt 60  
aggncgtntg ggctaaagggt ctctttgagc ccagtggctt tagtcacacc ttctttgctc 120  
tggncaggaa ggccctactt ttctntactc gtggaatctt ggaatcttaa agataaaaga 180

acctagaaaag aaaatcaaac ccacttcct tgngggcaga tggtaatatg ggactgagac 240  
agcaaacctg gggcttggan aggaccnanc tc 272

<210> 22  
<211> 270  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 40, 44, 46, 114  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700060207F7

<400> 22  
tacccctgca tgggataaccg tttctcgacc ccagtgcacn tggntnctgt catcccatga 60  
aattgcagca agggcagtct cttttgtggg aacagattaa ctcc tacaca tgangtagat 120  
tcaacacacctg ccagggaaagc agaagcatta cttaaatgtc ctgtgaaggc aaacatcaag 180  
tcaattcagc ttatctgaa gagtgccaaa ccatgaactc caaatgtcat tgtgtgaaac 240  
tgaacgatgg tcatttcatt ccgggtctgg 270

<210> 23  
<211> 250  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 45, 49-50  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700133105F6

<400> 23  
ccctcctgta tctgaaccca gcttctcagc tctgagatga gtgcnggann ggcttcccaa 60  
cctatgtca ataccacagg cagcctgcag gagggagaaa tggtaaaat gttccatggg 120  
aaatgtctca gaatcgctc ccccgaaatct cctgctaagc tttactgtc ctatggagt 180  
atcatggtcc tcagtgttagc tgttagttgct ctttctgttg ctttgtcagt aaaaatgaca 240  
ccacagatct 250

<210> 24  
<211> 226  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700138117F6

<400> 24  
gaggattcac tcacatttgc ttcccgctgg ccatgagtga gctgccctt ctgagtcag 60  
agggagccag agggcctcac aacaacagag ggtctcagag ctccctggag gaaggctcag 120  
ttacaggctc agaggctcgg cacagcttag gtgtcctgaa tgtgtccttc agcgtcagca 180  
accgtgtcgg gccctggtgg aacatcaa at catgccagca gaagtg 226

<210> 25  
<211> 265  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 85  
<223> a or g or c or t, unknown, or other  
  
<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700268788F6

<400> 25  
cgagggtgct cccaggcggc tgcactggct cgaggaggta gcaggaggag ctccgcgcag 60  
gaacaaacct ggaggcaaac caganggagg caatgttga atgactgtaa gaagaccaga 120  
cagtaaaaat gtcagccctc aactggaagc ctttggtgtc cgagggctg cctccatcac 180  
cgcagaatgt ggtacatttc caattgattt gactaagact cggttcaga ttcaaggcca 240  
gacaaaatgt gccaagttcc gagag 265

<210> 26  
<211> 257  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 58, 99  
<223> a or g or c or t, unknown, or other  
  
<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700270924F6

<400> 26  
ctgggatccc cagggcta at gggcatcctg ttcttcgcagc agggcactgt gagaagnct 60  
ctcaccgtga ccaaggttct ctgagtgtcc agccaaccna ggctcaccag ctccctccag 120  
ctaccgcccc tccatcaggc cagctgccaa ccccaggctg aacaccaacc ccagctatga 180  
gctcctggag gcatgactcc ctcaggccca gcagctccga tccccctccca gtagttatca 240  
ttggcaatgg ccctcgg 257

<210> 27  
<211> 244  
<212> DNA  
<213> Rattus norvegicus

<220>

<221> unsure  
<222> 2, 8, 50, 56, 63, 76, 177, 219, 233, 240  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700303722F6

<400> 27  
gnccgggnca ccaagggttt tctgcctcta cttagaaaagt ccaaggggan gctggntaac 60  
gttnacgacca tggganccat gattccattt cagatgatgg ccgcctacgc ctgcacgaag 120  
gcagctataa gcatgttctc agcccgatcat caggcaagag ctccaaatg gggagtnaaa 180  
gcgggaccat cattctggag cttcaaaccac acatcgatng ctcacaggac agntggatn 240  
aaag 244

<210> 28  
<211> 263  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700502047F6

<400> 28  
ggaaatgact ggtctgaagc ggcttggcaag cctgagcagt caggtactgc ggctctactg 60  
gcactgcctg agggtccaag gactgtggtg attctcatgg aggaccctga gatttctgca 120  
atctgatcaag tgtcaaatgc cactggattt cgtctgagac tcttgcctta gaggatggcc 180  
aaagggtctcc tgatgaccta tgccctttgg ctttggggc cctgttggac tacaccacct 240  
gtatctggga agggacagcc atc 263

<210> 29  
<211> 259  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 52, 137, 167, 189, 235, 254, 256  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700502260F6

<400> 29  
ggactacccc caccatgccc gtgtcattga gactctgatt gtccactatg gnctggctt 60  
tgaggaggag ccagaagaag cagctggcaag ccaaggggg gcgtccgcca gtgtgccag 120  
ctggagactg ctgaggncat tgtcttcccc cagcaggagg aggccggncga tggaaaccga 180  
gaatcccang tgcataatg actcagactc agagctggaa gaggcttctg acctnttcg 240  
cctcgacgc cacncnctc 259

<210> 30  
<211> 260

<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 140, 158, 161, 170-171, 177, 191, 193, 199, 203, 206-207, 212  
<222> 217-218, 220, 224, 226, 231, 238, 243, 246, 248, 253, 255, 258  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700142213F6

<400> 30  
cacagggttt tccccaaagag cacgcctt tacttcaggg aattctcgga aactttcagg 60  
attcgggggg ctccctggg aggagacgag agaggattag aagcggacat ccacggcttc 120  
ttgtgatgac cacgccttn gtcttgcta gaactctntg ngacttccn nggtganttc 180  
taatcacgga ntncacgcna agnttnngga antacgnngn cccntnaaga naaacacntt 240  
ttnggngng ggnanaanct 260

<210> 31  
<211> 288  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700229555H1

<400> 31  
caggagtacc actcacaggc cacctggcag gaagagataa gcccccagcc cccgacatcc 60  
aggacgcccc gaacctgcca atgtgtgtatg ctatacctta ttacctcatc atgtgaaata 120  
gccaatcata tgtgaacatg tctatgtgcc tcgtttgaat ccaccaatcc ctgtaactat 180  
gcatctgctt ctgtacgcct gcttctgctt ccccaatccc tataaaagcc ccatgctgga 240  
gctgctggcc gcgcaagtcc tcctaagaga ctgtgtgccc gcagtacc 288

<210> 32  
<211> 258  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 53, 244  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700626839H1

<400> 32  
gggtgcgcgt ggagttgcgc atgcgccttc cggccgcgca gggcaaagggt ggnngcgc 60  
tgttgaatgg ttgggtgctg tgcaagagcg ttttctggct ttgggtggcg aaggcggct 120  
ggccgcgagg tgcagctgct ggtggggcagg tgtactaatg tctacagact atgagcttc 180  
agaatctctg gagagagtac aaagttctgc atgttatggt acctttaatc gggttcatac 240

attnnnnntgg cacagaat

258

<210> 33  
<211> 268  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 7, 11, 40, 53, 150  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700627890H1

<400> 33  
gtacaangag ngccggggct tgggtctagt tggaggggan gcagtggcca gtncaggcgt 60  
cagatgagag agtttagccga gtttagggca gctacttagga tgggggcagg aggagaagcg 120  
gggctaacta taaaagaagac tagatttcgn cacagtggtt atgttggaaagg cagcttcaa 180  
accgccttg tcaaacaaca cagggccagc agccttcaag accaggctat ccctgccgtc 240  
tgctggcatg ggggcacttg taccgtcc 268

<210> 34  
<211> 299  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700059105H1

<400> 34  
tcatcacaac ccaactgtgt ggccaaatcc agaggtgttt gacccttatac gatttgcacc 60  
agagtcttcc cgacacagcc actcattcct gcccttctca ggaggagcaa ggaactgcatt 120  
tggaaacag tttgccatga atgaactgaa ggtggccgtg gccctgaccc tgctccgctt 180  
ttagctgtgt ccagatcccc ccaggatccc aatccccata ccaagactcg tggtaagtc 240  
caagaatggg atctacctgc gtcctaaaaa gtcctaaaaa tcttgacagc acaagacag 299

<210> 35  
<211> 300  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700059610H1

<400> 35  
aaacaacacgtg actttcttgc gtgtgaggag tgccttttat gggAACAGCA tcatctacaa 60  
tatgtcctct gatggccgtt tgcccccgg ggcctgccag attgctcatg agcacacaga 120  
tggagtgtat aaaatgagga aggttcagct gcagaatgag gaagagcttc agaaggccag 180  
gaagaagagg cacttgatt tcctggacat cctgttgttt gccaaaatgg aggtatggaa 240  
gagcttgatc gatgaggacc tgcgtgcaga ggtggacaca ttcatgtttt agggtcatga 300



<210> 39  
<211> 79  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 39, 48, 66, 68, 72  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700608882H1

<400> 39  
cttaacgctc ctgccacgccc gcctccgccc gtgcaatgnc tctgtagneg gcgatctacg 60  
tagtntncc cngccccgt 79

<210> 40  
<211> 248  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 8, 21, 24, 35, 40, 52, 78, 90, 104, 137, 169, 209, 220  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700483988H1

<400> 40  
tttagtttca atcatggagg nctntcaggt gaggnctan ctcccaaagc cngcgctgag 60  
tctacaacctt ctcctaananag gtgttccggc cggcggttggg gtcnctgcgg aggccggctaa 120  
atcgccgcga gtttctncca tggttgcgcc cgctgtgttg cgcgctctnc gtaagaacaa 180  
gacccttcgc tatggagttc ccatgttgnt gctgggttgn agtgggttctt ttgggtcttcg 240  
cgaatttt 248

<210> 41  
<211> 352  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 28, 324, 337  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700501033H1

<400> 41  
ctctctgcct atgttctgag gttggagnct ttattacaga agctggtgca gaaaggagca 60

attgagaaaag aagtgtgaa tcaggcccgta gagaccaag tcattgctgg ggcaatccac 120  
aagtca gttc gaagagact tggactgcca gaaggtagcc ctgc cccagg cttattgcag 180  
ttgctgacac tgataaaaaga taaggaggca gcagaggaag agtcccttct tcaggccgaa 240  
ttagaaggac atttcacttg acccaagacc agcaaggctg tcatgagcag aacatgatgg 300  
aggagctcat agaagtgatc agcnccatccc ctttgggnctg ccaagtaatt gc 352

<210> 42  
<211> 233  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700137747H1

<400> 42  
gttctccat agcctcagac cccacatca gatcctcttgc tctacttggag gggcacgtga 60  
gcaaaaggc taaccatcta atcagcaagt tccagaagct gatggcagag gttggccact 120  
tcgaaccagt caaccagggtg gtggaatcgg tggctaatgt catcgagcc atgtgttttg 180  
ggaagaacctt ccccaaggaa agcgaggaga tgctcaacacct cgtgaagagc agc 233

<210> 43  
<211> 243  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 33  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700141708H1

<400> 43  
tggcagaaaa ggaaggccctg cagagcatca gangcccagc tagagggaca acacagagga 60  
gtaatttgc gacagacctg cagggatgga cctgcttca gcttcacac tggaaacctg 120  
gttcctcctg gcagtcgtcc tggctccttctt acggattt gggacccgca cacatggact 180  
tttcaagaaaa cagggattc ctggcccaa acctctgcct tttttggca ctgtgctgaa 240  
tta 243

<210> 44  
<211> 295  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 286  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature



<223> Incyte template ID No: 700325693H1

<400> 47

gtgccctcac gcagcttaat gtggccttt cccgggagca ggcccacaag gtctatgtcc 60  
agcaccttct gaagagagac agggAACACC tgtggaaagct gatccacgag ggcgggtccc 120  
acatctatgt gtgcggggat gctcgaaata tggccaaaga tgtgaaaac acattctatg 180  
acattgtggc tgagttcggg cccatggagc acaccCAGGC tgtggactat gttaaaaagc 240  
tgatgaccaa gggccgctac tcactagatg tgtggagcta ggagcttacc aaccccac 300  
ccctcg 307

<210> 48

<211> 300

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 21, 49, 199, 226

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700059571H1

<400> 48

ctgtgctct gagtgcaagc naggcattcc tccaagacac tgcgggtcng agcaggact 60  
gttcacgctg gtgcctgtg aactctggtg gaggtcagcc aacagctgt gtgtctgagt 120  
tgctgagagg agagagaatg gttgcactg agttttttt ccacgtgcca agtctggagg 180  
agctcgacaga agttttgcng aaggggctaa aggacaactt tgctcntgtc caggtctctg 240  
tggtcgactg cccagattta acaaaggagc catttacttt cccgtaaaag gcatctgtgg 300

<210> 49

<211> 314

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 13

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700060610H1

<400> 49

gcaagattga ggnggagaag gacaacctga agtctgagtt ccatctggag aacttggctg 60  
tctgtgggtc taacttgtt acggcaggca cccgagacaac cagcaccacc ctgagattcg 120  
ggctcctgtc ccttatgaag tatccagagg tgcaagccaa agttcatgag gaacttggacc 180  
gtgtgattgg acgccaccaa cccccccagca tgaaggacaa gatgaagctg ccttataccg 240  
atgctgtatt gcatgagatt caaaagataca tcactctcct tccttccagt ctgccccatg 300  
ctgtggtcca ggac 314

<210> 50

<211> 312  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 78  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700063128H1

<400> 50  
cggtcggta cggagagcgc aggttgtatc accaacatgg gggactctca cgaagacacc 60  
agtgccacca tgcctgangc cgtggctgaa gaagtgtctc tattcagcac gacggacatg 120  
gttctgtttt ctctcatcggt ggggtcctg acctactggt tcatcttttag aaagaagaaa 180  
gaagagatac cggagttcag caagatccaa acaacggccc caccgtcaa agagagcagc 240  
ttcgtggaaa agatgaagaa aacgggaagg aacattatcg tattctatgg ctcccagacg 300  
ggaaccgctg ag 312

<210> 51  
<211> 248  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 64-65  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700139656H1

<400> 51  
caaaaatacaa caaggagata aaagtcatac agtttgtgct gctggcttat tagctctgca 60  
tggnnngaggg gccatgttaa gttccaagg cattcagata taaatgtaat taagatgcca 120  
tgtttgcttg cagtaatgaa gtataatca gaaactgta aagtatgata aaaacagtga 180  
ttgttatgc acttatggaa gacaaagtga agtgatgtgg tttcttcaga acaggtgatg 240  
cactgagg 248

<210> 52  
<211> 115  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700141348H1

<400> 52  
gtaaaagatg tggccagag aacgttagaa catgcctgga gagaccta at gtgctcttgt 60  
tctgcaaacc catggcatt attccctct ccgctcaaga gctcatactg gaagc 115

<210> 53  
<211> 249  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700182318H1

<400> 53  
agagttccct tttgtcccc aacctgttagt tctaagttca acaaaaacagt catcaacaaa 60  
agtgcacggag gttccatca gtgtcagggc tctgtcaga cccagaatcc ctgttcccta 120  
tgtcatgttc cagcattgta tagcacggtt ccatgtcaca aacagaaaagg tcaggaacac 180  
tgaggctctgt gaatgtcaact gctgcagcga ggtcatgtca ctccctctgtc tactctgtca 240  
gtgtcttac 249

<210> 54  
<211> 296  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 22  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700225376H1

<400> 54  
agcaaggccta tttctgactg gnctgctgtc cagaatctag accactggca gtgggtgaca 60  
gccctgttgc ggtaatcgaa agtctcgtag caggctctgc tgtaagtctg gcctcttggc 120  
ctcacatctt ctttgggta tccttcccta tctccagctt cctcagctgg tcagggagat 180  
ttggtccaga actagaagcc ttaataatct gagcaggtaa gagaggagta aaatgtacag 240  
tcttggacat tgactaaagg gtcctgcaga ggatatcaag gtaagtggct tggagg 296

<210> 55  
<211> 169  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700225757H1

<400> 55  
gctttctggg caagtctatg ttgccttagc tgaccaggaa tttgctataat agactattct 60  
gtctcaaact cacagaaaatt ctccctgcctg tgccctctga gcgagcacca ggattaaagg 120  
cgtgaatcgc tgtccccgtc tttttcttt cttctttaa taacccact 169

<210> 56  
<211> 191  
<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 190

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700268788H1

<400> 56

cggaggtgct cccagggcggc tgcaactggct cgaggaggta gcaggaggag ctccgcgcga 60  
ggaaacaaacc tggaggcaaa ccagaaggag gcaatgtttg aatgactgtta agaagaccag 120  
acagtgaaaa tgtcagccct caactggaag ccctttgtgt acggagggct ggcctccatc 180  
aacccgcggan t 191

<210> 57

<211> 249

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 45, 118, 128, 163, 245

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700270924H1

<400> 57

tgggatcccc aggggctaattt gggcatcctt ttcttcgcgc aggnactgt gagaaagtct 60  
ctcacccgtga ccaagtttctt ctgagtgccc agccaacccca ggctcaccag ctccctcnag 120  
ctaccgcncg tccatcaggt caactgccaa ccccaggctg aanaccaaacc ccagctatga 180  
gtccctggag gcatgactcc ctcaggccca gcagctccga tccctcccaag tagtgatcat 240  
gggcnaggg 249

<210> 58

<211> 294

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 19, 32, 35, 104, 131, 188, 220

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700291661H1

<400> 58

actagcttctt ccagctccntt ctttcccgag angcngcagg gacctcgcc tccagcttac 60  
cgggcggatc gaagcagcgg tcggatggt actgctggc ttgtgcagt caggcggctc 120

```

ggtgtctcgaa naggcgcgtgg  ggtgac aggaggcaac ctgtttcca cgctgctcat 180
cgccctgcncc ttcaacgttta  ttgtcta cctgttccogn ctgcgactgg gccacatgg 240
ccagctgcccc gctggagccaa agtccggcc atatatttc tctccaattc cgtc 294

```

<210> 59  
<211> 304  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700301979H1

<400> 59  
gatataattaa tcaaaaagaa aggaccacaa ctcatgacct cccatcttcc catgcatttt 60  
ttctccaagt ctctcttcag ttgtatctt seqtgtatct atctcgtag aaatcccaga 120  
gatgttttttg ttctggttt attcaa ctcttgcgaa gaagccagac 180  
tcactgggaa cttatgttga ttctatatgg atcatggttt 240  
gagcacatcc gtgcctggct gccccatgcaa gaatggaca acttctgtt actgtactat 300  
gaag 304

<210> 60  
<211> 293  
<212> DNA  
<213> *Rattus norvegicus*

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700302770H1

<400> 60  
gtagccactc taacttagggg cgtgctgaga caagaccacc tcatttcctct gctgcttttc 60  
agacaggact gtctcgccga cccaccatga tccaggctgc actgttccctt ggctgtatct 120  
tactgtcctc ggtgaccggcc tttccatgga agactcagga tggtggcctg ccccatcago 180  
cagctggcac agaaactgag cctacacaac tgctctacag caagagtccct cctccgaccc 240  
ccagtacctg tcggAACCTC ctaagcatgg cgccccctgcc ccctgttagtc ctc 293

<210> 61  
<211> 174  
<212> DNA  
<213> *Rattus norvegicus*

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700303111H1

```
<400> 61
caacagcaaaa cgtgcgaccca actcttcggc tagtgaatct ctaagccgcg aagagtgcct 60
tgaaggtagct ttagtgtgaa gatgtcagaa agtaactcgg cagagggtag cgacacagaagc 120
qaggaggcagg tgctctggatc taaaatcatac qcccaaggcccc taaaaaacqca agat 174
```

<210> 62  
<211> 273

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 52, 139, 249

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700303390H1

<400> 62

ctcagatggc aggcatggca tgcgtggat cttgttcct gagacaaagg cngatgcaga 60  
ggccatgtga ataaatcatg aggggcccac agcaggccag caggccatag ctgaccatcat 120  
tctgaaagtg agagttgang agaccccagc tgggacagaa aaggtaccac gcctataacc 180  
atggcctaac cgagggccagc cagtgccagc ctccctgaaa gggacttcca gtccatccac 240  
aggcaccgna gaaccagcaa gacatagcca gcc 273

<210> 63

<211> 279

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte template ID No: 700303722H1

<400> 63

gtgactctga gtgttgagc aggtaacttc taccttgca cctctatcg aacaggtcca 60  
aagggtccaa aggagctggc aggacactca gacaagatcc actggctca ggtgtgccta 120  
gtcctggagt tcagaaagac ggaggcagct gaatgtggtg ctgaaccaac aacatctagc 180  
tacaagggga gccactcctc caccagcga ctgtgactgt tctcacaggt ctgaattcc 240  
tgttggatt cacaagatg ctttttattt ttaacttct 279

<210> 64

<211> 275

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte template ID No: 700306343H1

<400> 64

gagaaaaggcc accaccttagc taggtgagggt gtgccagcat ggtcctgggg gtctcactgt 60  
ccccagccct gggacgctgg ttccgcctatg caatccctt cgctatctt acgctttac 120  
ttctttatata cagtgtatgg ctcttccatg agtggccctt ttagttgcca gctcaaagaa 180  
ctcagcagtc cggcctgtgg gaactcaagc tctcttctcc ttctccagcc ctcacctctc 240  
tgcttcctgt cacctcaggt gttttacaag gctga 275

<210> 65

<211> 294

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte template ID No: 700306615H1

<400> 65

catccgtggg ctggctcacg ccattcgccct gttcctggag tatacagaca caagctatga 60  
ggacaagaag tacagcatgg gggatgctcc cgactatgac agaagccagt ggctgagtga 120  
gaagttcaaa ctggccctgg acttccccaa tctgcctac ttaattgtat ggtcacacaa 180  
gatcacccag agcaatgccca tcctgcgcta ccttggccgg aagcacaacc tttgtgggga 240  
gacagaggag gagaggattc gtgtggacgt tttggagaac caggctatgg acac 294

<210> 66

<211> 283

<212> DNA

<213> Rattus norvegicus

<220>

<221> unsure

<222> 2

<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature

<223> Incyte template ID No: 700330359H1

<400> 66

gnngcccaat ggcccctgag taaaaggtgg tcactgagag tcctaaggcc cgagtaggaa 60  
catgcggctt agagccagtg tcgtgaccca gaacgtccac tcttgcacag gtagatgagg 120  
aggtgttcgg gtggcccgag gcggtatccc gcctggctt cgccctagcc ttctgcac 180  
gcattggacat gaagccgctg gtggcctgg gactgcggc cccgacggcc cttccggct 240  
gtctctcctt ctggaaagct aaggcacagc ttgctcagag ctg 283

<210> 67

<211> 263

<212> DNA

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte template ID No: 700368974H1

<400> 67

tccttagctg gggcgaaaaa gggcagtctg ctactatctg aacgaattta atgtgggagt 60  
catgccttat acaacacagg aaattaatgt gtgatcta at gctgtatcta tgacttatta 120  
caatacagga attaaatgtg tgagccatgc cttcaaaaaca tgtcttagaaat ttctgaaatt 180  
ggccggaaatg caacaggat tgcttattta acctttcaaa tcactcattg tgacttagggc 240  
acatggtctt gcgcttgctta tga 263

<210> 68

<211> 269

<212> DNA

<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 129, 133, 153, 163, 165  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700373118H1

<400> 68  
gtttgaagca gggcgttaggg aaaagcggga ttaaagagta ctacctttat tagctccccc 60  
cctccagaat aggtgaaat ccctccccat gcccatttcc tgccacctgg ggtaaggatg 120  
tggcactgnc agnctgtcag cccactgact tttnagtcttc agntngcagt ctgggcaa 180  
accagcgagc tctgttgaac caagaccagg ctttcagagc atctgaacca ctgtggcctt 240  
ctctcctcag ctttcactgt ggcttttgc 269

<210> 69  
<211> 288  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 159  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700375521H1

<400> 69  
gcagccatgg atcgccccgg ggaacacctg tccgcgaggc cgccgcgttgg aaccgagagc 60  
ctgcgattcc tgcacgtcac agtgggctcc ctgctggcca gctatggctg gtacatccctc 120  
ttcagctcg tccttctcta cattgtcatc cagaagctnt ccctgcgact gagggctta 180  
aggcagagggc agctggacca agctgaggct gttctggagc ctgatgttgt tgttaagcga 240  
caagaggct tagcagctgc tcgtttgaga atgcaggaag atctgaat 288

<210> 70  
<211> 280  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 16, 84, 96, 112-114, 118, 171, 226, 228, 236, 240, 267  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700461546H1

<400> 70  
tgaacaaaact tctganaact tagttgacag tgtttgagt caactgaaaa aagcatgact 60  
tttggaatct ctgaatgcct tggnctctcag tattancact ctattgaatt tnnntctnat 120  
taaagtatgt agtttttag actttttcc tgacagtatt atgtaatttt ntggcgtggg 180

tagatggag tgtcgcttgt atgttaccat acagctgaca tgtatntttt gtctantctn 240  
attatcttag tagttcgtatg ctgtggnatg taccataacc 280

<210> 71  
<211> 271  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 25, 64, 70, 82, 102, 155  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700480077H1

<400> 71  
ctgacacctgac ccatgatgtta agggncgtta ggggagcatc accactgcaaa aggctgacta 60  
aggncgttn ggctaaagggt cnctttgaag cccagtgtct anagtcacac cttctttgct 120  
ctggggcccaag gaggcctact tcttctttt ctcnggaat cctggaatct taaagataaa 180  
agaacctaga aagaaaatca aaccacttt cttgtgggg cagatggtaa tatgggactg 240  
agaacagcaa acctgggtc ttggagagga g 271

<210> 72  
<211> 210  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 187  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700480949H1

<400> 72  
gggcagaggt ccagggaaata agggaggctt ctaccaatgtttaa tggtgcttga 60  
cagagatatt gtatggttct ctgagagctc ccctgaaaac cttacctcca accacacaag 120  
gttccctccc agagagcgct cgctggcag caaggacaca ctccccatatt gccaaagcata 180  
tcaagtnccc aaagattggc agaaaattcg 210

<210> 73  
<211> 256  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 70  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700483259H1

<400> 73  
gtgacgtaca tggaaaacaa agcctacggg gacaggctca agccgcagac agcagcaagt 60  
aaagcgctn cgccctgaa gcatggcagc tatcccttcc agcggctcg tcgtggctac 120  
ccatgactac tatcgcgta agtagcccc cggccagggct gcccccagg 180  
tctgtggctg acccgctcc cttcccagg acgtctggc tcctcggtcca gcaacagctc 240  
cgcggaagt gcagag 256

<210> 74  
<211> 259  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 219  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700483475H1

<400> 74  
ctctcgtaact tcggcaatgg ctggattccc accgtcatca cggccttgc cttgtacc 60  
tcccaggccc aagctggatg gtcataacat gattatggcc acctttctgt ctataagaaa 120  
tccatatgga accacattgt ccacaagttt gtcattggcc acttaaaggg tgcctccgccc 180  
aactggtggaa accatcgaca tttccagcac catgcgaanc caacatcttc cacaaggacc 240  
ccgacataaa gagcctgca 259

<210> 75  
<211> 264  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700498995H1

<400> 75  
gacagttagat gcccccaaag ctctagtaga tgatagtgtg gggctgtgt gggctcccta 60  
cctgtgtgt tcattcacag tgcaagttaa gggagcaggc gccactgcatt tccttgctg 120  
tgccctgagg gtgcttgctg ctttatatacg taacagtcaa ttaaggttca ttcaggaag 180  
agaaaaggga tggtttgag gggctcagaa aataggattc agtgtgttaac ataacaggta 240  
ggttgcggc acatgctgat atcc 264

<210> 76  
<211> 271  
<212> DNA  
<213> Rattus norvegicus

<220>

<221> unsure  
<222> 218, 228, 255, 270  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700504333H1

<400> 76  
gtttatTTTg acacagacat ggacaaAGCG atggaggcgct atgtctctat gcccAaggAA 60  
aaggGCTCCAG aacacATTCC CCTTCTCTTC ATTGCCTTCC catcaAGCAA ggATCCAACC 120  
tgggaggacc gattcccaga ccgatccaca atgactgtgc tggtaCCcAC ggcCTTgAA 180  
tggTTCgagg agtggcagga ggAGCCTAAG ggCAAGCNAA gtGTTGcNTT ggaACCCtCA 240  
aaaaaaACTTC CCGGNAACCC TTTATGGGm A 271

<210> 77  
<211> 167  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 11-12, 17, 21, 24, 48, 66, 72, 96, 128, 135, 162, 166  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700505040H1

<400> 77  
caacaatCTT nngtgnctg nctntctgga actgggcATC atcAGCTnAT gctGCCATAC 60  
gcCTGnTgGA gNGCCGTggG gtGAAGGTcG cCCGtnCCCT ggtGGGTAcc tTCATGTCAG 120  
cactaganAT gcgtngtGTT tCCCTTACTT tgatGTTGT gnatgna 167

<210> 78  
<211> 267  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 5, 18, 26, 39, 90, 92, 122-123, 132, 137, 145, 152, 160, 168, 171  
<222> 173-174, 186-187, 213  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700505423H1

<400> 78  
ggggNCTTCT gtGAGGCNCT gatacncATC gaggCTGTnA ttCAgCCAGg CCACATGAAG 60  
ccccAAGATg ggtGGCTTT CCTGTATGAn tnAGTACAGA tATATCCATg gCCGGGGAAT 120  
tnnGACTGGC angGTcncca gggancacca gncAGCTTn tcaAGAANTC ntNnGGTCC 180  
cttGGnNTCA cAGGAACCTA ttACCTTCA tGNGGTCTGG ggttCTGGAT ttAGGGTCTT 240  
tggGACAGTC ccAGTAGAA GcCTTGG 267

<210> 79  
<211> 267  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 14, 20, 22, 24, 81, 248, 253  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700510534H1

<400> 79  
tgaaaagcgta aggncttgcn tntnagaggc tctgttagtga gttctgtttg cctataaggg 60  
aagtggaca accgagacac ncgcacttct ttgcagggtt aaggagcctg ggaggagcag 120  
gcagccgctt gctttgagca tgctcagggtg gggctgtccg ccgcgtgtgg gaaggcaccc 180  
tgccagcaggg ctccctgccc cacctctcca ttgttagtagt gtccagatct cagaaacgca 240  
gcttgaancc agnttcaaag gtaccag 267

<210> 80  
<211> 291  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700513027H1

<400> 80  
gggagggtttt aaaggccata ttgccaacct caccgaaagg tttcaggaac ccgaggaagt 60  
gttaatgtac aactcaccac ttcacgccac ccgaggctga agttgacgtt gccttttaag 120  
cccttttaca tacactggcc atttcagaaa attctcaaca ataatgtctg cttcgagtt 180  
taagtcatgg tgtttttag aattgacttg aaatgaaaat atcacaaaatg gaatatatca 240  
gctggtgatc gagtgactga aacccccctg gtctgcgggtt gaccagttca g 291

<210> 81  
<211> 273  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700528082H1

<400> 81  
ataaaaagtga aaactgggc aagggcagggg gctggcgtg aaccgcttac tagataatgt 60  
tctctaaaaa ttggctctga aaaccctgtt tgttatcg ttttatgagt gctaaaaaat 120  
ggtgtgacca gggcatggtc actgtcattg gaacagcaac atgcttgcgt gtcacattgg 180  
aatggggaaa tgtgaagaaa gctggacatc aggctgcgg caccatttc tttgtatgaa 240  
agtgttgtgt acaaaaccccc cactaatcat ttt 273



tccccgggaac actaaaaggt agtcttactg tccaccaccc tacacctgtt ttcataagtt 240  
atgcacaaat gcgaacagct gagacagaga tggagaagtt cttcgttttg 290

<210> 85  
<211> 275  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 17  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607183H1

<400> 85  
cacagccccct accagcnac cctccataaac tgcaccaaga ggatctatcc aacacctccc 60  
tgagcaggag gagcctgaag actccaaggg aaagagtccct gaggaaccct ttccctgtgca 120  
gctggatcta accacaaaacc cacagggtga cacactggat gtctccttcc tctacctgga 180  
gcctgaggaa aagaaaactgg tggcctgcc ttccctggg aaggaacagc gctccctga 240  
gtgccccgggg cccgaaaaagc aaagaacccc ctgat 275

<210> 86  
<211> 285  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607235H1

<400> 86  
ctgaagaccc accatgtctc tgctgactac tgtactactt ctctgggtt tcattctggg 60  
cccagcaact gacacagcct gatatcaa ggaaggctcg gaaaacagtc ccttgcctcag 120  
gcccctggctt tctgccaatc cagtgcctcg gatcacacct ggcctgagga cattcctgct 180  
gtgccagggg acagtgcggg atgttagtctt catgtgagg cgggaaggag atgatggtt 240  
cctggcgata gtccaaacaga tgttttctg gagggagctg gaccc 285

<210> 87  
<211> 260  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 246-247  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607396H1

<400> 87  
ggccaccaag atggcggcgc ccagcggacg gtgcgcgagg ttcgtcagct gacttgtct 60  
cgagctgtg gcccgcaccc gtttctacct gtccccagtg accagagctc agtgaccagt 120  
ccttatagtc gaaagcaggg ttttactgc tgaggacctg gacccgctgg gaggcttgcc 180  
atggtaacag aacaggaggt agaggccata gggaaaaccc tagtgactc cacgcagccc 240  
ctgcannccc gttccgtgc 260

<210> 88  
<211> 181  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 27, 180  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607505H1

<400> 88  
caaagaaaaga aaacactcct ccgaggncgg cagcaaaggc agagaaaagat ctgcaggatg 60  
actttcatta cagcagggtt tatattttat ctttttgcc tccgttcta gtgaatgtat 120  
cactaagggtc ttcaaaagaca tcagcttca aggaggtgcc taagtactgt ttccacacan 180  
c 181

<210> 89  
<211> 280  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607713H1

<400> 89  
aattgagggg taaacatgtc tttgtgaaat atatgttctt ttacaatact ttgtactaat 60  
ttacgtggaa ttattatttg ttctcattt gagatattt ttcagcctca atggcccttc 120  
aggaactcct gaatcaggtt ggaggccttag ggagattcca gatcctttag atggttttg 180  
ttgtcttcac cagtttattt gtgttacctc atattataat agagaactta ctgcagccat 240  
tcccagtcat cgctgctggg ttcctatcct cgacaatgtc 280

<210> 90  
<211> 267  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 6, 77, 128, 143, 146  
<223> a or g or c or t, unknown, or other

<220>

<221> misc\_feature  
<223> Incyte template ID No: 700607873H1

<400> 90  
gaaganacac caaagctcac tcactttca ggttctcaact gaacctactg tatcggaagg 60  
gacttcacct cccagangtc cattttatg aagactgtt agacagctt ccagaaacta 120  
gaaccatntg gaagatagac ctnggngtat tcctgtcgg attatcttga ttacgttaat 180  
taattctgga tggactagg ctaaagtgtc atcatgattt tccattaaca aggtgcacag 240  
atgctacaaa tggctggag aaatcct 267

<210> 91  
<211> 258  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 11, 109  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700607972H1

<400> 91  
gaggctaagg ntgtgcctcc tggctctctt ggggcagtcg cccgcgcgcc aagactttc 60  
gctgacctca gcgtcccgt gctgcgcaag gaaggcgccc gccactgcng tctggacagc 120  
gtccgaaggc agcgagtctt ctggaggccg ccgttagtgca gaggagtcgg ttgtcacgtg 180  
acccaagggtt agaccatggc ttccaccaag ccgctgtctc gttctggga gtggggcaag 240  
aatatcgttt gcgtgggg 258

<210> 92  
<211> 276  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700608519H1

<400> 92  
caccacactg catctgcctt acatgccacc ttacaccatc atctatttcc cctcccgagg 60  
tgtggctca gcgtttcaa gatgaaggcc acaaggccata gtcactcctg gggAACCCCT 120  
caacactggc gctggaggga tggccaaag accacgttgg gggcaagacc gagacttggg 180  
gcgggactac aattgtggtt ggtggggcca ggactgaccc tttagcctcc ataggcagct 240  
acactgtctg tcactttcc tcctccctg tgccgg 276

<210> 93  
<211> 295  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure

<222> 35  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700608661H1

<400> 93  
cttagtcctt gtactctgag ggtaagcctc atcgntcagg atctattgct gctgcttc 60  
cagctggctt ctgttagagac tgacagagac aactttgcag gataaggtag ctatcaagat 120  
gctccatgg cgaaaggcac agatgctgca gatgtggag agctccttaa gaaaaataacct 180  
tcctgagtcc ttaaaaggttt atgggactgt ctccacatg aaccaggag ccccattcaa 240  
gctcaaggct ctggggaca agtggcctga tttaatacag tgggtgtcc tcctc 295

<210> 94  
<211> 293  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 16, 179  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700609074H1

<400> 94  
ggcgtggagt tggagnagag cgtcaggcgc ctccggaga agtttcatgg aaaagtgtcc 60  
cccaagaagg caggggctct tatgaggaag tttggcagcg accacactgg agttggcgc 120  
tctatcggt acgggctcaa gcagaaaagat ggacaggagc tgagcaacgca tttggacgnc 180  
caggaccac cagaggacat gaagcaggac caagatatcc aggcagtagc cacctctcg 240  
ttgccccctga cgcaagccaa tcttcgaatg ttccaaagag cccaagatgaa cct 293

<210> 95  
<211> 288  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 66, 138, 152, 236, 242, 252-253, 279  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700609967H1

<400> 95  
ctgcctcgca gccccgagcg cgcgcccttc cagctccccgc tccggcttcc ccaaccaggc 60  
ttattntgtc tcccgaccccg gtgcagaccc ctgacccggc ctccgccccag ctccgccaaa 120  
tgcgtactt tacttggtag gaggtggcgc angctccggg agggagaagg agcgatggct 180  
ctaatacgac cgaaagggtgt acaacatcag cgacttcagt tcgccccac ccgggnggct 240  
cncgggtcat cnccactag ctggtcagga tgccacggna tcctttgt 288

<210> 96  
<211> 164  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700625315H1

<400> 96  
gttagagcag ttacactgaa ccaaagtgac tgagttgta cagacggtaa tccgtaccaa 60  
gcacactcac tgtcctgatc tgaacaccca gcaaggttca tgtccgtgct aagttgcag 120  
cattgtgttc ttttgcatc ttttttact ttttattaaag gttc 164

<210> 97  
<211> 225  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 28, 59, 65, 73, 79, 84, 88  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700627089H1

<400> 97  
cgccggcgcgc accatgg nag agcttcagga ggtgcagatc actagaggng 60  
aagcnattgt tgnccaggntt ggcnttantg gttgagtcta tcactccagt gggttccctg 120  
tttgctctgg catcatactc catcatcttc ctcaagctt ttccttaccg ggtatgtcaat 180  
ctgtggtgc gcagcgaag ggtcaaggcc aaagctgtgt ctgca 225

<210> 98  
<211> 265  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 62, 264  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700629293H1

<400> 98  
atgaccttta acttttctaa aaatgtgaag ttttgtactt atatatatca gctaaagtat 60  
tn>tagcattc ttttagtgtac tttagttgat gccactttta gtgttttgt tgctttgtc 120  
tgatTTTtat gaatgttcat tttaagactc cttgttgaaa tgggacagtt tgcgttcttg 180  
ataagccccga gaagaggatt cccttgggtg ttgacctcct ctgcattatg tgcccaagca 240  
tctgaactgc aaccaaggcc ttnc 265

<210> 99  
<211> 95  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No: 700643961H1

<400> 99  
agtataacca gggccatct gaaagttgtt ctcttagccag ataagccact atgagcggta 60  
agcatcagtg ctaacgcaga aacttcctga gcagc 95

<210> 100  
<211> 307  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 8, 20, 21, 23, 47, 302  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700061625H1

<400> 100  
gggaaganac actggcttan nantgggttc tgagggaagc tctctgngtg gatacatatt 60  
tataggattt gggtgcaacg aactgtgtga ctatgttcaacttcaattcagg gagcaggcag 120  
aaaccagcag gtattagaag agatgttcta tatacagagt tctgaggcac tgcagattct 180  
gaagaattcc ctaaggaagc acctccctga gtccttaaag gttatggac tgtctccaca 240  
tgaaccaggg aaaccatttc aagctcaagg ctgtggtgga caagtggctg atttaatact 300  
gntgtta 307

<210> 101  
<211> 97  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 48, 90  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700062809H1

<400> 101  
ccctgagagt ctccaaaatg ttccgtctg ttataagaac cccattanct cactaaagct 60  
tgcaactggct gactcctgga cggggttccn gaatgtg 97

<210> 102  
<211> 214  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700139104H1

<400> 102  
gagatgttcc ctgtcatcga acagtatgga gacatggtaaaaatactt gaggcaagag 60  
aaaggcaaac ctgtccctgt gaaagaagtgtttgggtgcct acagcatgga tgtgatcacc 120  
agcacatcat ttggagtgaa ttttgattcc ctcacaacc cgaaggatcc ttttgtggag 180  
aaagccaaga agctcttaag aattgattttttt 214

<210> 103  
<211> 265  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700139953H1

<400> 103  
aggaagccct gcagagcatc agaggcccag ctagagggac aacacagagg agtaattgc 60  
tgacagacct gcagggatgg acctgtttc agctctcaca ctggaaacctt gggtccctcct 120  
ggcagtcgtc ctgggtgtcc tctacggatt tgggaccgc acacatggac ttttcaagaa 180  
acaggggatt cctggccca aacctctgcc ttttttggc actgtgctga attactatat 240  
gggttatgg aaattcgatg tggag 265

<210> 104  
<211> 306  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 2, 61, 68, 111, 139, 263, (296)...(298), 305  
<221> unsure  
<222>  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700289281H1

<400> 104  
ancagatccg ggcactggag agggagctgc aagcaggagc agtcaagagt gtggtcagaa 60  
ncaccgttag tgaccagca gggccggctt tttctaccat ggccggccaa ngctatggct 120  
attaccgcac tgtcatatnc acagccatgt tcggaggcta cagccttac tacttcaacc 180  
gcaaaacctt ctctttgtc atgccttcct tggtgatga gatcgctctg gacaaggacg 240  
atttgggggc tcacatcagcag canccagtcg gcagcctacg catcagcaag tttgtnnncg 300  
gggtnt 306

<210> 105  
<211> 183  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700303922H1

<400> 105  
ggcaggattt atccatatgt atatctgcc tttggaaatg gacccaggaa ctgcattggc 60  
atgaggtttg ctctcatgaa tatgaaaactc gctctacta aagttctgca aaacttctcc 120  
ttccagcctt gtaaggaaac acagataacct ctgaaattaa gcagacaagg acttctcaa 180  
cca 183

<210> 106  
<211> 290  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 56  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700305783H1

<400> 106  
ctaacagtga atacatagct gcgatcctgg aactcagtc ctcataatgt aaacgncaac 60  
gccagccctt cctgtacctg gacttcctgt attgcctcac tgctgatggg cggcgtttcc 120  
gcaaggcctg cgacgtggtg cacaacttca cagatgtgt catcagggag agacgcagca 180  
ccctcaatac ccagggcgtt gatgaattcc taaaggccag ggctaaagact aaaacttag 240  
actttattga tgttctcttg ctggccaagg atgagcatgg gaaggggctg 290

<210> 107  
<211> 177  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700329969H1

<400> 107  
gctatatacg gaggggaccc atgctgtgtc cttctgagat ctaacaggat taaccaatat 60  
gtaaaactaga ggaagtgggtt ggcctgcact gggcaagccc totaggactc catccaagaa 120  
agaccagttg gtgttgctct agaggcaaag aaaccataa ggagctggca gtaaaac 177

<210> 108  
<211> 188  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 114, 116  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700368493H1

<400> 108  
aaagcttccc aatctgtgtc ctcacttggg aggatggcat taggagccag ggctggccat 60  
gggtaccta ctttccccc tgggttatgc ccaggagaat ggagaaaaaa aaangntta 120  
aagaaaaaat atttaaatt tgatgctggc cttttcaat tgtattgagt aaaagtgttc 180  
aagttgtc 188

<210> 109  
<211> 255  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 156, 157  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700376694H1

<400> 109  
ttatgattat caattttaca taacattaat attatatcaa acctccttaa gaaaatgagt 60  
atggatgttc acagtatgtt tgattttat ctacaagaat gaatctgatt cagaatgctt 120  
ttcagctgac atacagagca ctaaatactt taagglnaac cataggtctg aatctcttaa 180  
gaattctcag tctctatggg atgttagggac gcattataaa tgcattaatc cttatagtca 240  
atcctgtgcc tagga 255

<210> 110  
<211> 284  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 16, 25, 54, 63, 68, 70, 80, 141, 154, 210, 275  
<221> unsure  
<222>  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700483803H1

<400> 110  
gaggaacgcc gccgcntcgc tcgnatcct acaccaatca ggaagctgct gtcnagccat 60  
ggnggganan gagaagccan ctcaagaggc tgacgtggaa cccatggtaa catcagggc 120

ctcagaagca gtgccaaagg ngctttctgg agancctcg aacatctctg atgttagatgc 180  
cttcaacttg ctcctggaga taaaactgan acgacggcgt gaggtcccaa cttccatgt 240  
actgtgaccc agctagtggc cgaggatggc agcanggtgt atgt 284

<210> 111  
<211> 258  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700502844H1

<400> 111  
ggcagacatttggatgccatc agtgactgaa gccattaagg agactcggt 60  
ggatggcag ctaagaggac agagagacaa gggctacgag gcagcaatctt 120  
ggtgttgcgt agatttgaga cgaagggttc ccatggcttc ttccatc cggcaggttcc 180  
aggagaggga ctatgaacag gtcgtggata tggttccatc gggaaatgaag gaacacatcc 240  
ccactgcctt ccggcact 258

<210> 112  
<211> 250  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700503415H1

<400> 112  
gtagctttcc ccttttgcgt gcacagaagt ctgtccatct gcaagcgctt tggaacacag 60  
actgccttggc gcccacccatc ttggggagac ctccctgcct cagctgtcg 120  
cattcaactaa agctccgtac gtcagattaa gcaagcgttg atgggttaca ttagagacaa 180  
ggccgcagaga taaggcctgt tgctgtttcg cagataatga tgagttttaa ttacccactg 240  
gtttgtatgg 250

<210> 113  
<211> 278  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 40, 44, (53)...(55), 60, 69, 72, 74, 159, 234  
<221> unsure  
<222>  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700514914H1

<400> 113  
ccactgcage gccccccccc aaagatggaa agattctgcn ttancttcat atnnnccttan 60

aacatttgna cncnattctt taactttaga atctcccta gagcctgttc ttctttaaac 120  
accctttatg ctggagtagg atgatggctg agtttcttna aaagagctt aatatagagt 180  
cacaaacatg agatagatgc ctgccgcca ctcttccac aaactcgaga accncttg 240  
gcacgcagct gccatggaag gaaatcctgg ggcttctt 278

<210> 114  
<211> 308  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700519169H2

<400> 114  
cccaatctt taagacttt cataatgatg ttaagaccag ggcagactat ctactggca 60  
tggactgcag ccaggagatg ccctgcccct cctggaccc gcacaccacc tctctgggga 120  
acttgacaaa ggtccctaag gctaagggag gtctccttcc tactaggtcc ctgactttga 180  
ctctgtgggtt ctctaggaac cgtgtgcaca cttgtctctg ttgttaaccac aaagggcagt 240  
agcaccttagc atgtcatgtc ctgccccggc tgcttgccctc ccacccaccc aggattctct 300  
gggctggc 308

<210> 115  
<211> 124  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 34, 53, 87, 109, 122  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700535905H1

<400> 115  
cctacctctg gtggaaaaag tacatcaactc aggngcagct ggtccagttt gtngctgaca 60  
atcatccaga ccagctgcgg ggtcatntgg ccgtgctctt tccctctcng gtggctgtac 120  
tncc 124

<210> 116  
<211> 262  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 148  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700593984H1

<400> 116  
gggaaaagtgt gttctgaggg ccctgtgggg ccaaggggga ccagcctcac attccacacg 60  
tgcgccactc tgcttggagc ctatttattt tgtatttattt tgaacagagt tatgtcctaa 120  
ctattttat agatttgtt aattaatncc ctgtcatttt caagttcatt ttttttattc 180  
atatttatgt tcatggttga ttgtacctcc tgtcaccagc tggtgggca. ggggagacaa 240  
ggtagaaaagt ggccacacagag tg 262

<210> 117  
<211> 267  
<212> DNA  
<213> Rattus norvegicus

<220>  
<221> unsure  
<222> 36, 37, 223  
<223> a or g or c or t, unknown, or other

<220>  
<221> misc\_feature  
<223> Incyte template ID No.: 700607844H1

<400> 117  
ctcgatcgtc cagacccac ggctgtcacca tgctgnncctc tgatcagggaa ctcactgttc 60  
gggagcgtgg agacgtggcc ttggcaggtt ttaagcacccg ggggcaagga agacgtctcc 120  
tatgagggaaa gagcctgcga agggggaaag tttgctactg tggaaagtgc agacaaacct 180  
gtggatgagg ctctccggga agcaatgccc aagatcatga agncgtgggg ggcaccaatg 240  
acaaaggagt cgccatggga atgacag 267



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> :  <b>C12Q 1/68</b>		A3	(11) International Publication Number: <b>WO 00/12760</b>  (43) International Publication Date: <b>9 March 2000 (09.03.00)</b>
(21) International Application Number: <b>PCT/US99/19768</b>		(US). ZWEIGER, Gary, B. [US/US]; 2700 Del Medio Court, #124, Mountain View, CA 95023 (US). PANZER, Scott, R. [US/US]; 965 East El Camino, #621, Sunnyvale, CA 94087 (US). SEILHAMER, Jeffrey, J. [US/US]; 12555 La Cresta, Los Altos, CA 94022 (US).	
(22) International Filing Date: <b>27 August 1999 (27.08.99)</b>		(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).	
(30) Priority Data:  09/141,825                    28 August 1998 (28.08.98)            US 09/172,711                    13 October 1998 (13.10.98)          US 09/172,108                    13 October 1998 (13.10.98)          US		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications  US                            09/141,825 (CIP) Filed on                      28 August 1998 (28.08.98) US                            09/172,711 (CIP) Filed on                      13 October 1998 (13.10.98) US                            09/172,108 (CIP) Filed on                      13 October 1998 (13.10.98)		(71) Applicant ( <i>for all designated States except US</i> ): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).	
(72) Inventors; and (75) Inventors/Applicants ( <i>for US only</i> ): CUNNINGHAM, Mary, Jane [US/US]; 1204 Manet Drive, Sunnyvale, CA 94087		Published <i>With international search report.</i>  (88) Date of publication of the international search report: <b>3 August 2000 (03.08.00)</b>	
(54) Title: TOXICOLOGICAL RESPONSE MARKERS			
(57) Abstract  The present invention relates to a composition comprising a plurality of nucleic acid molecules. The composition can be used as hybridizable array elements in a microarray. The present invention also relates to methods for screening compounds and therapeutics for metabolic responses indicative of a toxic compound.			

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Larvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 99/19768

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL [Online] emb1 heidelberg AC#H33194, 30 September 1995 (1995-09-30) LEE H.N. ET AL.,: "EST108960 rat PC-12 cells" XP002130308 abstract &amp; LEE N.H. ET AL.,: "Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment" PROC. NATL. ACAD. SCI. USA, vol. 92, - August 1995 (1995-08) pages 8303-8307, XP002047703 the whole document</p> <p>---</p> <p>-/-</p>	6
X		6

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

14 February 2000

19.05.2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Müller, F

1

## INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/19768

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 13877 A (LYNX THERAPEUTICS INC ;MARTIN DAVID W (US)) 17 April 1997 (1997-04-17) see whole doc. esp. claims and p.4, line 12 ff. ---	1-5,7,8
Y	WO 95 20681 A (INCYTE PHARMA INC) 3 August 1995 (1995-08-03) see whole doc. esp. claims and p.40, line 3 ff. ---	1-5,7,8
Y	SCHENA M ET AL: "PARALLEL HUMAN GENOME ANALYSIS: MICROARRAY-BASED EXPRESSION MONITORING OF 1000 GENES" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, US, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 93, no. 20, 1 October 1996 (1996-10-01), pages 10614-10619, XP002912238 ISSN: 0027-8424 the whole document -----	1-5,7,8

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 99/19768

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-8 (partially)

#### Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-8 (partly)

Nucleic acid sequence (seq id 1), fragments thereof, methods for detecting and diagnosing metabolic effects of compounds using microarrays comprising seq id 1 or any fragments thereof and methods for screening a library of molecules for specific binding to seq id 1.

2. Claims: 1-8 (partly)

Nucleic acid sequence (seq id 2), fragments thereof, methods for detecting and diagnosing metabolic effects of compounds using microarrays comprising seq id 2 or any fragments thereof and methods for screening a library of molecules for specific binding to seq id 2.

inventions 3-117

...ibidem for each nucleic acid sequence or a fragment thereof separately seq id 3-117

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

Interr	nat Application No
PCT/US 99/19768	

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9713877	A 17-04-1997	AU 712929	B	18-11-1999
		AU 4277896	A	06-05-1996
		AU 6102096	A	30-12-1996
		AU 7717596	A	30-04-1997
		CN 1193357	A	16-09-1998
		CZ 9700866	A	17-09-1997
		CZ 9703926	A	17-06-1998
		EP 0793718	A	10-09-1997
		EP 0832287	A	01-04-1998
		EP 0931165	A	28-07-1999
		FI 971473	A	04-06-1997
		HU 9900910	A	28-07-1999
		JP 11507528	T	06-07-1999
		JP 10507357	T	21-07-1998
		NO 971644	A	02-06-1997
		NO 975744	A	05-02-1998
		PL 324000	A	27-04-1998
		WO 9641011	A	19-12-1996
<hr/>				
WO 9520681	A 03-08-1995	US 5840484	A	24-11-1998
		AU 688465	B	12-03-1998
		AU 1694695	A	15-08-1995
		BG 100751	A	31-07-1997
		BR 9506657	A	16-09-1997
		CA 2182217	A	03-08-1995
		CN 1145098	A	12-03-1997
		CZ 9602189	A	14-05-1997
		EP 0748390	A	18-12-1996
		FI 962987	A	26-09-1996
		JP 9503921	T	22-04-1997
		LV 11696	B	20-08-1997
		NO 963151	A	27-09-1996
		PL 315687	A	25-11-1996
		HU 75550	A	28-05-1997
<hr/>				

**THIS PAGE BLANK (USPTO)**